

THIS WEEK

EDITORIALS

SOUND Tune into the harmony between music and science **p.262**

WORLD VIEW Vaccine makers can learn from space industry **p.263**



WHAT DRESS? Original colour-shifters show their secrets **p.265**

A clean, green science machine

As the world warms and technology improves, researchers and institutions should look at their carbon footprints and question whether they really need to travel to academic conferences.

Every time the United Nations climate negotiations get under way, media stories appear about the carbon emissions generated as thousands of government officials, environmentalists and scientists fly in from around the world. Similar questions have been raised about major environmental-science conferences, such as the Fall Meeting of the American Geophysical Union (AGU), which last year drew an astounding 24,000 people. But rarely does this discussion move beyond the obvious. It can indeed seem a bit disingenuous that people who are trying to understand and protect the planet will engage in such a high-carbon activity as aviation, but what is to be done?

Researchers at the Tyndall Centre for Climate Change Research in Norwich, UK, have published a thoughtful working paper (see go.nature.com/zdztid) that lays out a practical methodology to tackle these questions, from the standpoint of both individual academics and institutions. The bottom line is that aviation emissions are rising, and that academics in the industrialized world are responsible for more than their fair share, thanks to the countless conferences, the international nature of science and the need to do fieldwork in far-flung locales. If academics are to fall into step, they must curb their aviation emissions in concert with the rest of the world.

Academics in all disciplines — not just climate experts — should read the paper, not least because it goes beyond posing moral questions and delves into solutions. Simply purchasing carbon offsets will not balance things out, the authors argue, because the market for reliable offsets may be limited and the net effect could be to legitimize unnecessary travel. If scientists want to bolster their credibility on the subject of global warming, the authors say, then they must harness the power of the Internet and reduce the time they spend in the air. Indeed, the authors propose that a greater reliance on online conferencing and social media will not just reduce emissions, it will make science more inclusive. The conferences that do need to take place could be more centrally located for the target audience. And to keep track of it all, the paper provides an accounting system that allows institutions to log “hours in motion” and calculate the resulting emissions.

The arguments the authors put forward are powerful, and there are good reasons to pursue their recommendations. In fact, some quarters have already come to similar conclusions and are taking appropriate action: the AGU’s leaders, for instance, have started to look for ways to hold parallel conferences and online dialogues that will allow more academics, often from developing countries, to engage in its meetings. There is real value in face-to-face interaction, and we must not forget that. But the Tyndall Centre is right to point out that senior researchers probably do not need to fly halfway around the globe simply to present a paper at a conference.

In some ways, the working paper opens the door to questions that are even harder to answer. Do scientists have a responsibility to stop eating meat, given what we know about the greenhouse-gas intensity of beef production and to a lesser extent that of pork and chicken?

Should we expect them to park their cars and take the bus or train instead? The fact is that these are personal choices that academics, like everybody else, must grapple with.

Nor does the buck stop with individuals. Institutions, academic or otherwise, have a large influence on the daily lives that their employees lead. It might seem extreme for universities to force vegetarian fare on their students (although many institutions now have a meat-free day), but bosses could go a long way towards reducing traffic jams and carbon emissions by encouraging their employees to telecommute. And in the long run, does it make sense for an institution or company to purchase cheap real estate in the suburbs and then force its employees to bear the cost of commuting?

There are plenty of solutions to be had at many levels, and all options may need to be exercised to stave off the worst impacts of global warming. But lots of seemingly small actions can have an effect: last week, the International Energy Agency reported that global carbon emissions stalled in 2014, even as the global economy grew by 3%, apparently thanks in part to efforts by China and industrialized nations to boost energy efficiency and adopt renewable sources.

Irrespective of what academics do, it seems likely that steaks, cars and planes will not only persist but will increase in number as the population booms and becomes more affluent. Ultimately, the world must identify better ways to raise beef and move people around. And scientists have a key role in making that happen, even if it means hopping on a flight to the next United Nations climate summit. ■

“Lots of seemingly small actions can have an effect.”

Hollow humour

The public’s distorted attitude towards mental-health conditions hampers their treatment.

What is so funny about obsessive-compulsive disorder (OCD)? Paul Cefalu, a professor of English at Lafayette College in Easton, Pennsylvania, asked that question in a 2009 article that analysed why OCD is so frequently, and unfairly, represented in the media with “humor and levity” (P. Cefalu *PMLA* <http://doi.org/c2mtt4>; 2009). Unlike other mental-health disorders, such as depression and schizophrenia, those who suffer from obsessions and compulsions, Cefalu wrote, “can always be counted on to make us laugh”.

The laughs are thinner on the ground in *This Room*, an autobiographical one-woman play about OCD, written and performed by Laura Jane Dean. In a review on page 289, Emily Holmes calls it

“an affecting, sobering account of a life shaped by, yet transcending, a mental-health condition and treatment”.

The portrayal of mental-health conditions (or, to be less semantically guarded, mental illnesses) in the media and popular culture has a significant influence on the way that many people view both the conditions and those who have them. Next to lawyers and police officers, physicians are among the most frequently fictionalized professionals, and psychiatrists feature heavily. Despite the well-quoted statistic that one in four people will experience a mental illness at some point in their lives, the medical reality is something more easily viewed at a distance, peeking through the silver screen at the misery of somebody else. When that picture is distorted, the result is fear and mistrust, not just of the mental-health professionals but also of their patients.

So why do Hollywood and the rest of popular culture often get psychiatric medicine so wrong? Why do we still talk of the psychiatrist's couch when most people who see a psychiatrist never sit on one?

Too many accounts show only the illness, or at least a crude stereotype of it, not the person. And, crucially, they do not show how the person can often improve with the right treatment. Miracle cures are as rare in mental as in physical medicine, but many (perhaps most) people with psychiatric problems who receive the right kind of help can start to feel better. It is important to acknowledge and portray this.

A study published in February in the journal *Social Science & Medicine* (E. E McGinty *et al.* *Soc. Sci. Med.* <http://doi.org/2v4>; 2015), for example, suggested that stories about people who were successfully treated for depression, schizophrenia and drug addiction reduced negative attitudes towards mental illness in those who heard them. Tales of the same problems left untreated produced more willingness to discriminate against sufferers. This discrimination and prejudice is known to act as a barrier to people's recovery, because they can experience

social rejection when word of their condition gets around.

Perhaps here is a way to tackle the stigma of mental illness. It is not enough for scientists, journalists and campaigners to ‘raise awareness’ by highlighting the reality of such conditions, and by publicly criticizing those who misuse the terms and language of psychiatric medicine. It is not enough to point out that phrases such as ‘schizophrenic foreign policy’ and ‘a little bit OCD’ are offensive because they misrepresent

“The way to break the destructive cycle, to treat the OCD, is to resist the ritual.”

and trivialize genuine and serious suffering. It is not enough to emphasize the biological basis of mental illness. To truly change public attitudes, the message must go out more often that this suffering can be alleviated.

In doing so, a powerful cultural myth must be challenged: that mental illness is a gift and comes with benefits. From the supposed enhanced creativity and meaningful visions of those with schizophrenia to the claimed cognitive wizardry and insight of people with autism, mental-health conditions are too often presented as just another way of seeing the world. This sense of instant karma might soften the blow to audiences, to some patients even, but it feeds the damaging impression that psychiatrists are out to rob special people of their gifts and unique potential. It is hard enough for people with mental-health problems to seek help, without their fearing those who are best placed to provide it.

What is so funny about OCD? The answer, Cefalu concluded, is the incongruity of the condition: the harder a sufferer tries to help themselves with comforting rituals, the worse their torment becomes. The way to break the destructive cycle, to treat the OCD, is to resist the ritual, to stop playing to the crowd. That was what Laura Jane Dean managed to do. That was her treatment. That is her story. And stories do not have to be funny to have a happy ending. ■

Strike a chord

The latest episode of the Nature Audiofile podcast looks at how music inspires science.

“**M**athematics and music! The most glaring possible opposites of human thought! And yet connected, mutually sustained!” Thus enthused German scientist Hermann von Helmholtz in an 1857 essay on harmony in music.

Today, mathematicians, ecologists and physicists search for harmony alongside musicians, if not always explicitly. But the link used to be even stronger.

A typical academic curriculum circa 1600 — and for centuries before — blended music and (what we would call) science to a degree rarely seen in today's undergraduate syllabi. The four subjects of the ‘quadrivium’ were arithmetic, geometry, astronomy and music.

Today's scientific training programmes tend to leave music out. They have changed in other ways too, of course, and for the better. Although some universities offer degrees in mathematics and music, or physics and music, modern scientists more often miss the chance to seek musical connections.

“Many people think music is a charming accompaniment to thought,” says musician and scientist Peter Pesic of St John's College in Santa Fe, New Mexico, in the latest episode of *Nature's* sound-science podcast, *Audiofile* (go.nature.com/xhluk3). “But developments in music,” he notes, “influence other aspects of human thought.”

The podcast — one of a series — contains plenty of musical stories from the history of science. Take the astronomer Johannes Kepler, who was preoccupied with the motions of the planets. He was desperate to find harmony that he felt sure existed in the way the Universe was set up. In 1619, he produced a giant volume called *Harmonices Mundi*,

or *The Harmony of the World*. In one particularly musical moment, he expresses planetary motions in musical notation — the orbits of Mercury, Venus and their neighbours spun out into crotchets and quavers. You can hear the resulting song on the podcast. Without such explicitly musical thinking, it is possible that Kepler may never have arrived at his third law of planetary motion — the relationship between a planet's distance from the Sun and the time it takes to orbit.

Then there is the tantalizing suggestion that Galileo Galilei's musical father might have influenced the way his son thought about science. In the late 1580s, Vincenzo Galilei carried out an experiment on the sounds made by strings held at different tensions. Vincenzo's home-made experiments could well have instilled in his son the idea of looking at a physical system to produce a hypothesis, rather than retrofitting the one to the other.

Musical analogies continue to help scientists to make sense of tricky concepts. String theorists speak of tiny vibrating strings instead of point-like particles. The comparison with stringed instruments is easy to see; the strings represent elementary particles in the same way that the strings on a guitar make different notes.

Musical inspiration can often remain buried under the surface of scientific work; there might not seem to be an immediate harmony between, say, a genome-wide association study and one of Erik Satie's études. But if music remains apart from the sciences, there is much we might lose. Anecdotal, neuroscientists talk of musical theory aiding the analysis of brain patterns. And a theory of strings is more intelligible to some than a theory of tiny, massless, vibrating subatomic squiggles.

Looking too hard for harmony might be misleading if the real picture turns out to be more discordant. But, as explained in the podcast by Jim Bennett, emeritus director of the Museum of the History of Science in

Oxford, UK, music has already provided a great template for interpreting our surroundings: “The insight, which plausibly came from music, that the world has a mathematical blueprint is fundamental to science.” ■

➔ **NATURE.COM**
To comment online,
click on Editorials at:
go.nature.com/xhunqv



Share the risks of Ebola vaccine development

Ebola vaccines have little in the way of commercial markets, so the risks should be shared between governments and industry, says Seth Berkley.

There are hundreds of infectious diseases out there that people could catch. More than 300 such conditions were discovered in the second half of the twentieth century alone. And how many of these diseases can scientists and clinicians protect against with a licenced vaccine? Fewer than 30.

Those are not always the biggest killers, or the most terrifying. Vaccine development is driven not by the risk that a pathogen poses to people, but by the economic pay-off. Given the difficulty of the science involved, how much money will it take to develop the vaccine? And given the size of the market, how much money can we make by selling it?

That helps to explain why, more than a year on from the first confirmed cases of the ongoing Ebola outbreak in West Africa, no vaccine is available, even though work started towards one more than a decade ago. Phase III trials for two vaccines have now been launched in Liberia and Guinea, and we have great hope for them, assuming that there are still enough cases developing to test the vaccines for efficacy. But for the more than 10,000 people who have lost their lives, and countless others who have suffered and will continue to suffer, these trials have come too late.

Our inability to protect people against Ebola is part of what makes the disease so frightening. In most cases, it is not what a disease is capable of that scares us, but that we can do so little about it.

But why is this the case for Ebola? We have known about the disease since 1976, and the first vaccine candidate was developed more than a decade ago. Ebola is not hypervariable like influenza or HIV, constantly changing and finding new ways to evade our immune systems, so we have had ample time to develop a vaccine or effective treatment during any one of the previous 23 outbreaks. Why were we caught by surprise this time?

The short answer is that we were not, but that the development of a vaccine was considered too financially risky. With a disease such as Ebola, which kills ferociously but occurs sporadically and usually in remote areas, there is simply no commercial market. Who would buy it? Outbreaks usually involve only a couple of hundred cases and occur every few years in poor rural communities in Africa. This leaves little in the way of incentives for manufacturers to invest the hundreds of millions of dollars it takes to develop a vaccine and get it clinically approved.

It is childish to blame the drug industry for failing to develop an Ebola vaccine — a product with no market. Instead, governments, public funders and private donors should be stepping up and investing.

We must work on a strategy that allows meaningful quantities of proven vaccines to be quickly produced and distributed when an outbreak occurs — of Ebola or other infectious diseases.

A first step is to identify the biggest threats, and

that demands better disease surveillance. More and better-equipped laboratories, as well as trained epidemiologists, in developing countries would improve our ability to quickly detect and investigate outbreaks of commonly occurring diseases, as well new threats.

The vast amount of data produced by this kind of surveillance network would have an added bonus. With the right smart data-mining algorithms, the information could be used to radically increase our understanding of how pathogens travel and mutate, and then how our immune systems respond to these changes.

When an outbreak occurs and vaccines are needed, it would help significantly to have vectors ready to deliver them. With the right investment, these vectors, typically a harmless virus or bacterium, could be prepared and tested in advance. Crucially, they could be pressed into service to tackle a range of diseases. Four of the five Ebola vaccines currently going through clinical trials use vectors developed and tested for HIV.

Such generic vectors would, in effect, modularize the vaccine development process — conducting much of the safety testing and ironing out manufacturing processes for different vectors ready for the addition of a 'payload' antigen. By developing such mechanisms in advance, and pre-testing them for safety and dose, we can save significant amounts of money and time by having stockpiles frozen and ready for use or efficacy testing as soon as an outbreak occurs.

This is similar to the way in which technology developed using public funds through NASA has reduced the cost of placing scientific probes, telescopes and satellites into space. Same rocket,

different payload.

It demands a different attitude to disease control. We need to stop waiting until we see evidence of a disease becoming a global threat before we treat it like one. Vaccine development is expensive, but the United States currently spends at least US\$11 billion a year to keep fleets of nuclear-armed submarines patrolling the oceans to protect people from a threat that will almost certainly never happen. That is 60 times more than the World Health Organization puts into global disease preparedness.

Governments and donors need to invest in public-health capability, and they need to take on more of the risk of investing in vaccine development. We must view vaccines as the ultimate deterrent: make sure they are there, and pray that we never have to use them. ■

Seth Berkley is chief executive of Gavi, the Vaccine Alliance, based in Geneva, Switzerland. This article is based on a TED talk delivered in Vancouver, Canada, on 18 March.
e-mail: sberkley@gavi.org

WE NEED TO
STOP WAITING
UNTIL WE SEE
EVIDENCE OF A
DISEASE BECOMING A
GLOBAL THREAT
BEFORE WE TREAT IT
LIKE ONE.

➔ **NATURE.COM**
Discuss this article
online at:
go.nature.com/8uxluy

RESEARCH HIGHLIGHTS

Selections from the
scientific literature

BIOGEOCHEMISTRY

Methane's great Arctic escape

Methane is moving from thawing Arctic soils into lakes and could be released into the atmosphere. Methane is a potent greenhouse gas, so this mechanism might exacerbate future Arctic warming.

A team led by Adina Paytan of the University of California, Santa Cruz, found that levels of methane were higher in soils around Toolik Lake, Alaska, than in the lake water itself. Geochemical measurements suggested that the methane is transported from the soil's active layer, which freezes and thaws every year, into the lake and then into the atmosphere.

If that pattern holds true for other northern lakes, soil could have a bigger role in sending methane into Arctic air than previously thought.

Proc. Natl Acad. Sci. USA <http://doi.org/25h> (2015)

NEUROSCIENCE

Nanoparticles turn on neurons

Gold nanoparticles can be attached to neurons and used to stimulate the cells, without introducing any genes.

Current 'optogenetic' methods use light to excite specific neurons, but genes must first be inserted into the cells to make them sensitive to light. To develop an alternative method, Francisco Bezanilla at the University of Chicago in Illinois, David Pepperberg at the University of Illinois at Chicago and their colleagues used molecules including antibodies to attach 20-nanometre-wide gold spheres to three different ion channels on the surface of cultured neurons. When

the researchers flashed a millisecond pulse of light, the gold heated up, causing most of the neurons to fire. The same thing happened when they injected the nanoparticles into a specific region of a mouse brain slice.

Neuron <http://doi.org/2sj> (2015)

MATERIALS

Liquid metal motor moves by itself

A tiny drop of liquid metal can propel itself for more than an hour without external help.

Millimetre-scale motors

could find uses as sensors, pumps and drug carriers, but they often require external drivers such as electric fields. Jing Liu and his colleagues at Tsinghua University in Beijing created a 60-microlitre liquid-metal motor that drove itself at around 5 centimetres per second by 'eating' aluminium.

The team applied flakes of aluminium to droplets of an alloy of gallium and indium. A chemical reaction between the aluminium, the alloy and a surrounding electrolyte propelled the metal beads around a Petri dish or through zig-zag and U-shaped channels.

The authors say that the work is a step towards creating a self-powered soft robot that can change shape according to its environment.

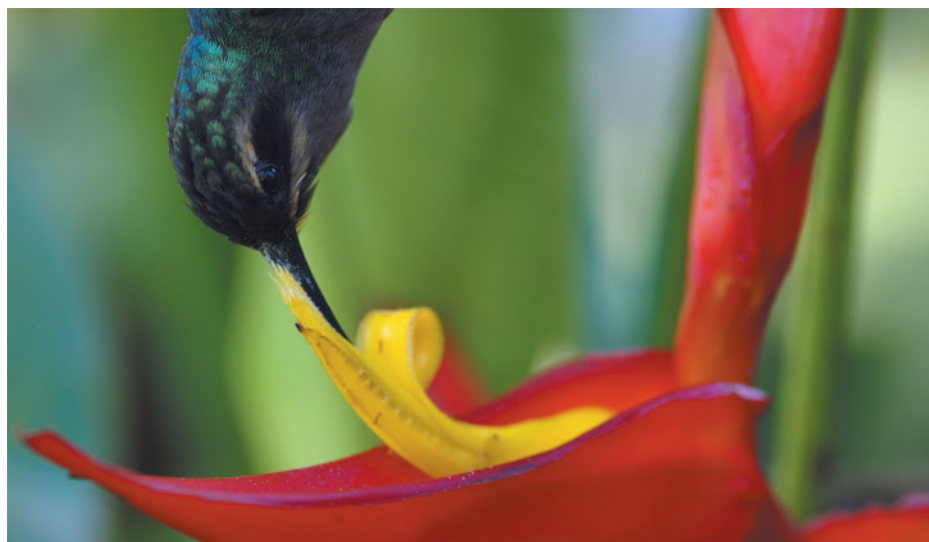
Adv. Mater. <http://doi.org/f26cb6> (2015)

NANOMATERIALS

Self-cleaning paint works in oil

A coating that can be easily applied to various surfaces repels water and dirt, even when exposed to oil.

Other water-repellent paints stop working in oil and are



MATT BETTS

ECOLOGY

Flowers choose the best pollinators

A tropical flower can turn on reproduction after it has been visited by a high-quality pollinator.

Matthew Betts of Oregon State University in Corvallis and his colleagues focused on the plant *Heliconia tortuosa* (pictured) and collected 148 of its pollinators, comprising six hummingbird species and one species of butterfly. The animals were cleaned of any pollen and introduced to aviaries containing flowers that had been hand-pollinated. The plants showed signs of successful reproduction

only after their nectar had been drunk by hummingbird species with long curved beaks, such as the green hermit (*Phaethornis guy*; pictured). Hummingbirds without the specialized bills and butterflies took in less nectar and failed to trigger reproduction.

Preferred birds also have the widest ranges, suggesting that the plants are boosting their chances of receiving pollen from distant flowers with more genetic diversity than nearby plants.

Proc. Natl Acad. Sci. USA <http://doi.org/2sf> (2015)

easily scratched. To overcome this, Ivan Parkin at University College London and his colleagues covered titanium dioxide nanoparticles with a hydrophobic polymer and suspended the particles in ethanol.

They sprayed or painted the suspension onto hard surfaces such as glass and steel, and dipped soft fabric materials into it. The coating repelled water and dirt, and did so even after being exposed to oil. Bonding the coating to surfaces using commercial adhesives made the film resistant to scratches from a knife and sandpaper.

The material could be useful in industrial applications that involve harsh and oily environments, the authors say. *Science* 347, 1132–1135 (2015)

CHEMISTRY

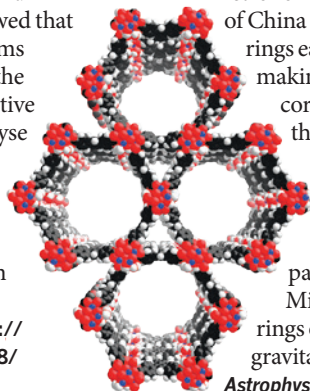
Metal framework zaps nerve agents

A crystalline compound catalyses the destruction of a nerve agent much faster than other clean-up chemicals do.

Omar Farha and Joseph Hupp at Northwestern University in Evanston, Illinois, and their colleagues studied a metal–organic framework (MOF) — a porous network of metal nodes linked by organic groups. They found that their zirconium-containing MOF (pictured) broke down half of a simulant of the chemical warfare agent DMNP in 15 minutes. Breakdown of 50% of the nerve agent GD took just 3 minutes.

Experiments and calculations showed that this MOF performs quickly because the zirconium ion active sites, which catalyse the breakdown reactions, are more easily reached by the nerve agents than in other MOFs.

Nature Mater. <http://dx.doi.org/10.1038/nmat4238> (2015)



MARINE MICROBIOLOGY

Microbes lurk deep below the sea

Microbial life may exist far deeper in the ocean floor than is often assumed.

Steven D'Hondt at the University of Rhode Island in Narragansett and his colleagues sampled sediments across the southern Pacific Ocean. They found that oxygen, and microbes that require it, permeated depths of up to 75 metres below the sea floor — more than double previous estimates.

The team found that oxygen penetrates the entire sediment column where the sediment accumulates slowly in a shallow layer. On this basis, the authors estimate that microbes that use oxygen may exist at low, but measurable, amounts throughout sediment in around 15–44% of the Pacific and in 9–37% of the global sea floor. *Nature Geosci.* <http://dx.doi.org/10.1038/ngeo2387> (2015)

ASTRONOMY

Milky Way has corrugated rings

The Milky Way's stars sprawl outwards in a series of concentric ripples, hinting that it might extend farther into space than was thought.

Data from the Sloan Digital Sky Survey confirm a previously known ring of stars at about 9,000 parsecs from the Sun. They also show another ring about 14,000 parsecs from the Sun, says a team led by Yan Xu of the National Astronomical Observatories of China in Beijing. These rings each form a ripple, making our Galaxy corrugated rather than flat.

The ripples may have formed when a dwarf galaxy passed through the Milky Way, creating rings of stars with its gravitational pull.

Astrophys. J. 801, 105 (2015)

SOCIAL SELECTION

Popular articles on social media

'Science fandom' can hurt science

Research stories that go viral on social media can bring science to a wider audience. But there is a downside to this 'science fandom', argues writer Ben Thomas in an essay on the Medium website that triggered discussion online (see go.nature.com/k9vwqj). Much of what gets shared lacks the nuance and uncertainty of science — a gloss that Thomas dubs "scienceyness". He writes that sharing the latest science headlines without any critical thought or fact-checking, whether by scientists or non-scientists, is contributing to an "onslaught of misinformation". Some on social media thought the blame was misplaced. Picking on the consumer who may not have science training "is a little unfair", says Lindsay Waldrop, a mathematical biologist at the University of North Carolina in Chapel Hill, who commented on the article on Twitter. Others suggested an upside to

scienceyness. Paul Coxon, a materials scientist at the University of Cambridge, UK, tweeted: "Scienceyness" isn't bad. It's often a way for excluded groups to get involved."

NATURE.COM
For more on popular papers:
go.nature.com/pcmqxc



BIOPHYSICS

Chameleons tune cells to change hue

Chameleons change colour by tuning nanoscopic structures in their skin cells to reflect different wavelengths of light.

Michel Milinkovitch and his colleagues at the University of Geneva, Switzerland, studied skin cells of the panther chameleon (*Furcifer pardalis*) from Madagascar. They found that the lizards have two layers of specialized cells called iridophores. Each layer contains light-reflecting guanine nanocrystals. By altering the spacing between the crystals in the upper layer, the cells shift from reflecting blue light to reflecting yellow or red wavelengths, which interact with the chameleon's yellow pigments. This produces a change in colour from green (pictured, left)

to yellow–orange (right).

The deeper layer consists of cells that reflect a broad set of wavelengths, particularly those in the near-infrared range.

The first layer of cells allows the animals to quickly switch between camouflage and an ostentatious display to attract mates or expel a rival male, whereas the second layer provides thermal protection. *Nature Commun.* 6, 6368 (2015)

CORRECTION

In the Research Highlight 'X-rays reveal virus innards' (*Nature* 519, 132–133; 2015), the image was described as showing the virus. In fact, it shows the X-ray diffraction patterns of the virus.

NATURE.COM
For the latest research published by Nature visit:
www.nature.com/latestresearch

Cooperative insertion of CO₂ in diamine-appended metal-organic frameworks

Thomas M. McDonald¹, Jarad A. Mason¹, Xueqian Kong^{2,3}, Eric D. Bloch¹, David Gygi¹, Alessandro Dani⁴, Valentina Crocellà⁴, Filippo Giordanino⁴, Samuel O. Odoh⁵, Walter S. Drisdell⁶, Bess Vlasisavljević², Allison L. Dzubak⁵, Roberta Poloni^{7,8}, Sondre K. Schnell^{2,9}, Nora Planas⁵, Kyuho Lee^{2,10}, Tod Pascal¹⁰, Liwen F. Wan¹⁰, David Prendergast¹⁰, Jeffrey B. Neaton^{10,11,12}, Berend Smit^{2,6,13}, Jeffrey B. Kortright⁶, Laura Gagliardi⁵, Silvia Bordiga⁴, Jeffrey A. Reimer^{2,6} & Jeffrey R. Long^{1,6}

The process of carbon capture and sequestration has been proposed as a method of mitigating the build-up of greenhouse gases in the atmosphere. If implemented, the cost of electricity generated by a fossil fuel-burning power plant would rise substantially, owing to the expense of removing CO₂ from the effluent stream. There is therefore an urgent need for more efficient gas separation technologies, such as those potentially offered by advanced solid adsorbents. Here we show that diamine-appended metal-organic frameworks can behave as ‘phase-change’ adsorbents, with unusual step-shaped CO₂ adsorption isotherms that shift markedly with temperature. Results from spectroscopic, diffraction and computational studies show that the origin of the sharp adsorption step is an unprecedented cooperative process in which, above a metal-dependent threshold pressure, CO₂ molecules insert into metal-amine bonds, inducing a reorganization of the amines into well-ordered chains of ammonium carbamate. As a consequence, large CO₂ separation capacities can be achieved with small temperature swings, and regeneration energies appreciably lower than achievable with state-of-the-art aqueous amine solutions become feasible. The results provide a mechanistic framework for designing highly efficient adsorbents for removing CO₂ from various gas mixtures, and yield insights into the conservation of Mg²⁺ within the ribulose-1,5-bisphosphate carboxylase/oxygenase family of enzymes.

Exceeding 13 gigatonnes (Gt) annually¹, carbon dioxide generated from the combustion of fossil fuels for the production of heat and electricity is a major contributor to climate change and ocean acidification^{2,3}. Implementation of carbon capture and sequestration technologies has been proposed as a means of enabling the continued use of fossil fuels in the short term, while renewable energy sources gradually replace our existing infrastructure⁴. The removal of CO₂ from low-pressure flue gas mixtures is currently effected by aqueous amine solutions that are highly selective for acid gases⁵. As a result of the large energy penalty for desorbing CO₂ from such liquids, solid adsorbents with appreciably lower heat capacities are frequently proposed as promising alternatives^{6,7}. In particular, as a result of their high surface areas and tunable pore chemistry, the separation capabilities of certain metal-organic frameworks have been shown to meet or exceed those achievable by zeolite or carbon adsorbents^{8–10}.

Recently, the attachment of alkyldiamines to coordinatively unsaturated metal sites lining the pores of selected metal-organic frameworks has been demonstrated as a simple methodology for increasing low-pressure CO₂ adsorption selectivity and capacity^{11–14}. Most notably, functionalization of Mg₂(dobpdc) (dobpdc^{4–} = 4,4'-dioxidobiphenyl-3,3'-dicarboxylate), an expanded variant of the well-studied metal-organic framework Mg₂(dobdc) (dobdc^{4–} = 2,5-dioxidobenzene-1,4-dicarboxylate)^{15–18}, with *N,N'*-dimethylethylenediamine (mmen) generated an adsorbent with exceptional CO₂ capacity under flue gas conditions and unusual, unexplained step-shaped adsorption isotherms¹³. Here we

elucidate the unprecedented mechanism giving rise to these step-shaped isotherms and demonstrate that replacing Mg²⁺ with other divalent metal ions enables the position of the CO₂ adsorption step to be manipulated in accord with the metal-amine bond strength. As we will show, the resulting mmen-M₂(dobpdc) (M = Mg, Mn, Fe, Co, Zn) compounds, here designated ‘phase-change’ adsorbents, can have highly desirable characteristics that make them superior to other solid or liquid sorbents for the efficient capture of CO₂.

Figure 1 illustrates the extraordinary advantages associated with using an adsorbent exhibiting step-shaped isotherms in a temperature swing adsorption process in comparison with the Langmuir-type isotherms observed for most microporous adsorbents. For carbon capture applications, a gas mixture containing CO₂ at low pressure (*P*_{ads}) and low temperature (*T*_{low}) is contacted with the adsorbent, which selectively adsorbs a large amount of CO₂. The adsorbent is heated to liberate pure CO₂ with a partial pressure of *P*_{des}, and is then reused for subsequent adsorption-desorption cycles. For a classical adsorbent (Fig. 1a), including all previous amine-based sorbents, the steepness of the isotherm gradually diminishes as the temperature increases, necessitating a high desorption temperature to achieve a large working capacity for a separation. In contrast, for a phase-change adsorbent of the type investigated here (Fig. 1b), the position of the isotherm step shifts markedly to higher pressures as the temperature increases, such that a large working capacity can be achieved with only a small increase in temperature. For an efficient carbon capture process, one would ideally create a phase-change

¹Department of Chemistry, University of California, Berkeley, California 94720, USA. ²Department of Chemical and Biological Engineering, University of California, Berkeley, California 94720, USA.

³Department of Chemistry, Zhejiang University, Hangzhou 310027, China. ⁴Chemistry Department, NIS and INSTM Centre of Reference, University of Turin, Via Quarello 15, I-10135 Torino, Italy.

⁵Department of Chemistry, Chemical Theory Center and Supercomputing Institute, University of Minnesota, Minneapolis, Minnesota 55455, USA. ⁶Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA. ⁷Université Grenoble Alpes, Science et Ingénierie des Matériaux et Procédés (SIMAP), F-38000 Grenoble, France. ⁸Centre National de la Recherche Scientifique, SIMAP, F-38000, Grenoble, France. ⁹Department of Chemistry, Norwegian University of Science and Technology, Høgskoleringen 5, 7491 Trondheim, Norway. ¹⁰Molecular Foundry, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, California 94720, USA. ¹¹Department of Physics, University of California, Berkeley, California 94720, USA. ¹²Kavli Energy Nanosciences Institute, University of California, Berkeley, California 94720, USA. ¹³Institut des Sciences et Ingénierie Chimiques, Valais, École Polytechnique Fédérale de Lausanne (EPFL), Rue de l'Industrie 17, CH-1950 Sion, Switzerland.

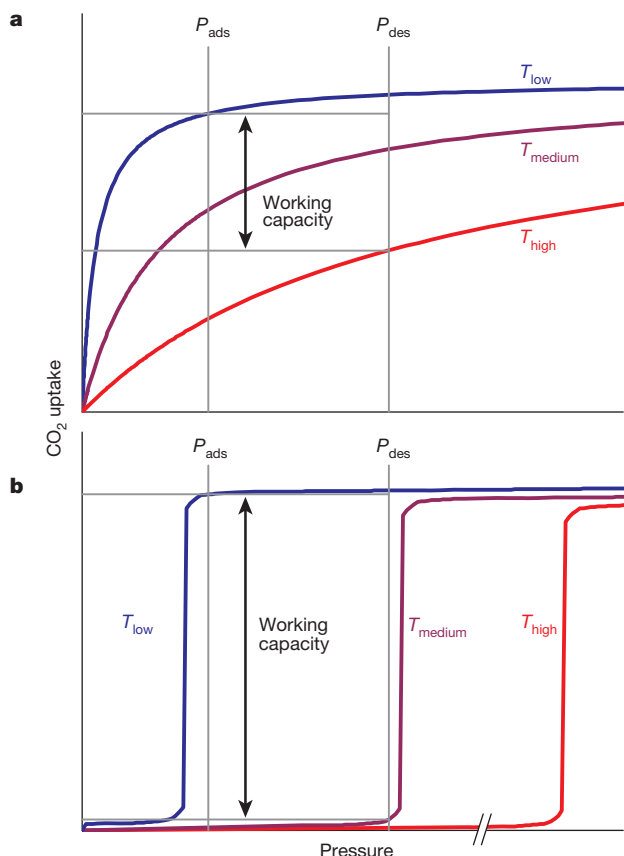


Figure 1 | Idealized CO₂ adsorption isotherms. Variation in the idealized adsorption isotherm behaviour with temperature for a classical microporous adsorbent (a), showing the usual Langmuir-type isotherm shape, compared with that of a phase-change adsorbent (b), showing step-shaped (sometimes referred to as ‘S-shaped’) isotherms. The double-headed black arrow indicates the working capacity (that is, the amount of gas removed) for a separation performed using a temperature swing adsorption process in which selective adsorption occurs at P_{ads} and T_{low} and desorption is performed at P_{des} and T_{high} (a) or T_{medium} (b).

adsorbent with a large vertical step positioned below the partial pressure of CO₂ in the flue gas.

Cooperative insertion of CO₂ into metal–amine bonds

Spectroscopic and diffraction measurements were undertaken to determine the mechanism of CO₂ uptake leading to a steep adsorption step for adsorbents such as mmen-Mg₂(dobpdc). In particular, powder X-ray diffraction studies, which were performed on the isostructural compound mmen-Mn₂(dobpdc) owing to the greater crystallinity of its base framework, provided detailed structural information on how CO₂ binds within the channels of the material. Diffraction data collected at 100 K before and after exposure of a sample to 5 mbar of CO₂ showed the unit cell volume contracting by just 1.112(8)%, but revealed large changes in the relative intensity of selected diffraction peaks (Fig. 2a). Complete structural models were developed for both data sets using the simulated annealing method, as implemented in TOPAS-Academic¹⁹, followed by Rietveld refinement against the data (Fig. 3, Extended Data Fig. 1 and Supplementary Tables 1–4). Before exposure to CO₂, the mmen molecules were bound through one amine group to the Mn²⁺ sites with a Mn–N distance of 2.29(6) Å, whereas the other amine lay exposed on the surface of the framework (Fig. 3c). Counter to our initial assumption that the uncoordinated amine groups would serve to bind CO₂ (refs 13, 20), CO₂ adsorption instead occurred by means of full insertion into the Mn–N bond, resulting in a carbamate with one O atom bound to Mn at a distance of 2.10(2) Å (Fig. 3d). The second O atom of the carbamate

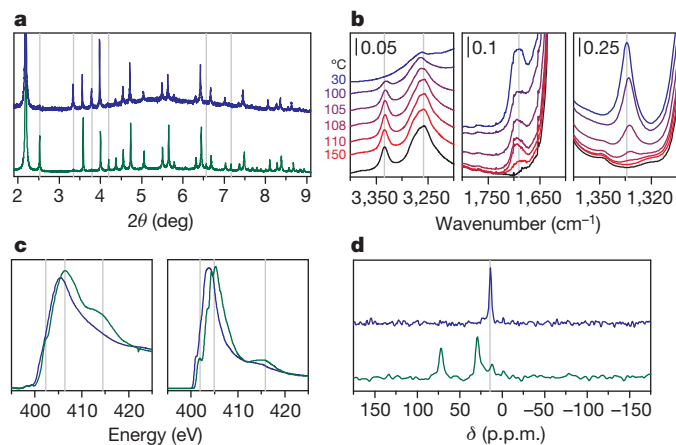


Figure 2 | Experimental characterization of the adsorption mechanism.

a, Large intensity differences are apparent in the powder X-ray diffraction patterns (collected at 100 K) on exposure of mmen-Mn₂(dobpdc) (blue) to 5 mbar CO₂ (green). b, Infrared spectra on dosing an activated sample of mmen-Mg₂(dobpdc) (black) with CO₂ and cooling from 150 °C to 30 °C (red to blue) under 5% CO₂ in N₂. The three different regions show bands corresponding to N–H (left), C–O (centre) and C–N (right) stretching vibrations. Spectra in the left panel are artificially offset by 0.05 a.u. to aid in visualization. Those in the other two panels are not offset; there CO₂ adsorption is responsible for the increase in the spectral baseline due to molecular charge delocalization of the ammonium carbamate chains. c, Experimental (left) and computational (right) NEXAFS spectra of mmen-Mg₂(dobpdc) at the N K-edge, before (blue) and after (green) CO₂ adsorption; the three major spectral changes are reproduced. d, Solid-state ¹⁵N NMR spectra for mmen-Mg₂(dobpdc) before (blue) and after (green) exposure to CO₂ at 25 °C.

had a close interaction of 2.61(9) Å with the N atom of a neighbouring mmen, resulting in chains of ammonium carbamate running along the crystallographic *c* axis of the structure (Fig. 3e). The observed ammonium carbamate N···O distance was similar to the distance of 2.66–2.72 Å in a single crystal of pure mmen-CO₂ (methyl(2-(methylammonio)ethyl)carbamate)²¹. This well-ordered chain structure was maintained at 295 K, as determined from a full Rietveld refinement against data collected at this temperature. Thus, the adsorption of CO₂ at ambient temperatures is associated with a structural transition to form an extended chain structure held together by ion pairing between the metal-bound carbamate units and the outstretched ammonium group of a neighbouring mmen molecule.

The foregoing structural information enabled the formulation of a detailed mechanism for the adsorption of CO₂ in phase-change adsorbents of the type mmen-M₂(dobpdc). As shown in Fig. 4, the uncoordinated amine of a mmen molecule acts as a strong base to remove the acidic proton from the metal-bound amine of a neighbouring mmen molecule. Deprotonation occurs only in the presence of CO₂, such that simultaneous nucleophilic addition of CO₂ results in the formation of a carbamate with an associated ammonium counteranion. At suitable temperatures and pressures, rearrangement of the carbamate is possible such that the M–N bond is broken and a M–O bond is formed. Critically, the ion-pairing interaction causes the mmen molecule to stretch, destabilizing the M–N bond and facilitating insertion at the next metal site. This cooperative effect will propagate until a complete one-dimensional ammonium carbamate chain has formed. Indeed, it is this cooperativity that leads to the sudden uptake of a large amount of CO₂ and a steep vertical step in the adsorption isotherm.

Infrared spectroscopy measurements performed on mmen-Mg₂(dobpdc) fully support the proposed mechanism. As shown in Fig. 2b, changes to the infrared spectrum were apparent when a sample of the compound was cooled isobarically from 150 °C to 30 °C at 1 °C min^{−1} under flowing 5% CO₂ in a N₂ atmosphere. At high temperatures, two distinct N–H vibrations arose at 3,258 and 3,334 cm^{−1}, which were also present in the spectrum of mmen-Mg₂(dobpdc) in the

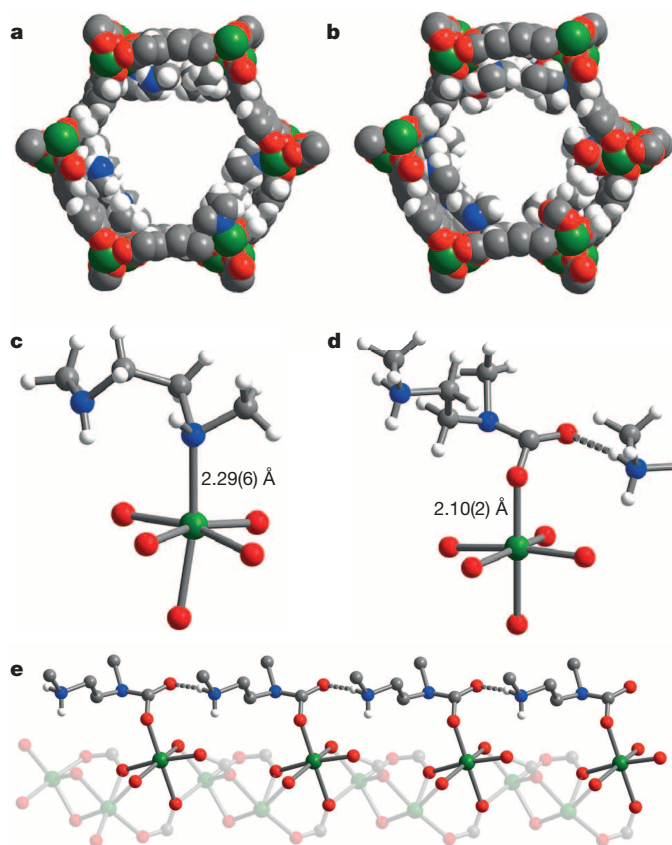


Figure 3 | Powder X-ray diffraction structures of mmen-Mn₂(dobpdc). **a, b**, Space-filling models of the solid-state structures of mmen-Mn₂(dobpdc) (**a**) and CO₂-mmen-Mn₂(dobpdc) (**b**) at 100 K. **c, d**, Portions of the crystal structures for mmen-Mn₂(dobpdc) before (**c**) and after (**d**) CO₂ adsorption, as determined from powder X-ray diffraction data. **e**, A portion of the crystal structure for the final configuration of CO₂ adsorbed within mmen-Mn₂(dobpdc), depicting the formation of an ammonium carbamate chain along the pore surface. Green, grey, red, blue and white spheres represent Mn, C, O, N and H atoms, respectively; some H atoms are omitted for clarity.

absence of CO₂ and could be attributed to the coordinated and uncoordinated ends of mmen, respectively. On cooling, both of these N–H resonances disappeared, indicating changes to both amines of mmen, while a new, extremely broad N–H band characteristic of ammonium formation appeared. From the weak but clearly discernible C=O vibration at 1,690 cm^{−1}, carbamate formation between mmen and CO₂ occurred under all conditions, even at high temperatures. However, an additional sharp band at 1,334 cm^{−1}, corresponding to the C–N vibrational mode of a carbamate, was observed only on cooling below 110 °C. The delayed onset of this easily recognizable band, which is diagnostic of a phase-change adsorbent of the type investigated here, is attributable to changes in the resonance configuration of carbamate that occur on coordination of one of its O atoms. The normalized intensities of the C–N band and a second band at 658 cm^{−1} versus temperature demonstrate that their formation was directly related to the sharp step in the gravimetric adsorption isobar measured under identical experimental conditions. From the infrared spectra it is clear that although a small amount of CO₂ can be adsorbed by means of ammonium carbamate formation between pairs of adjacent amines, it is specifically the adsorption of CO₂ to form ammonium carbamate chains that endows these materials with their step-change adsorption properties (see Extended Data Fig. 2 for additional infrared spectra).

To better understand the stepwise pathway by which the amines initially adsorb CO₂, density functional theory (DFT) calculations were paired with *in situ* near-edge X-ray absorption fine structure (NEXAFS)

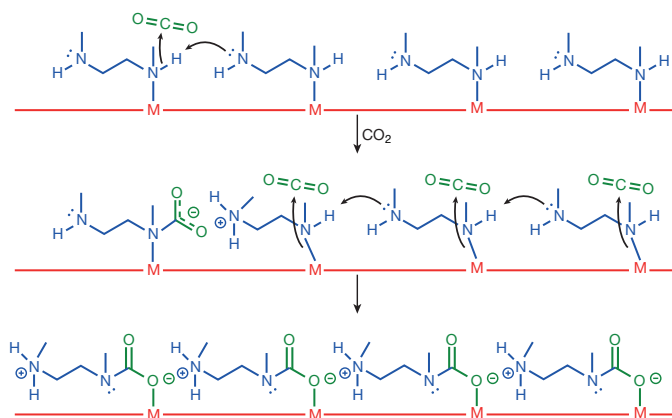


Figure 4 | A cooperative insertion mechanism for CO₂ adsorption. Depiction of the mechanism for CO₂ adsorption at four neighbouring M–mmen sites within an infinite one-dimensional chain of such sites running along the crystallographic *c* axis of a mmen-M₂(dobpdc) compound. Simultaneous proton transfer and nucleophilic attack of N on a CO₂ molecule forms an ammonium carbamate species that destabilizes the amine coordinated at the next metal site, initiating the cooperative adsorption of CO₂ by a chain reaction.

measurements of the nitrogen K-edge of mmen-Mg₂(dobpdc) collected under increasing CO₂ pressure (Fig. 2c and Extended Data Fig. 3)²², and all observed spectral changes were accurately reproduced by computed spectra. From the NEXAFS spectra, the new pre-edge peak at 402.3 eV arose solely from the carbamate nitrogen and was a clear signature of carbamate insertion into the metal–nitrogen bond. As in infrared spectroscopy, this feature is attributable to resonance of the nitrogen lone pair into the π system of the carbamate after the breaking of the coordinate bond using the same electron pair with the Mg metal centre. A second new, broad feature between 411 and 419 eV also arose solely from the carbamate nitrogen and was a signature of the new N–C bond formed on adsorption of CO₂. This feature appeared before insertion and was general to both terminal-bound and inserted carbamate moieties. Finally, the ~ 1 eV blueshift of the main edge peak at 405.4 eV was characteristic of ammonium formation.

Solid-state NMR spectra indicated that CO₂ adsorption affected the manner in which diamines were coordinated to the metal sites of the framework (Fig. 2d). On exposure of mmen-Mg₂(dobpdc) to CO₂, ¹⁵N chemical shifts consistent with ammonium and carbamate were observed at 31 and 72 p.p.m., respectively. Yet only a single ¹⁵N resonance was apparent for mmen-Mg₂(dobpdc) in the absence of CO₂. This indicates that the coordinated and uncoordinated ends of the mmen molecules were capable of interconverting on the timescale of the NMR experiment, although, as discussed above, they were distinguishable on the much faster timescale of infrared spectroscopy. Despite being labile, the amines were stable to evacuation under vacuum at high temperatures. This unexpected lability seems to allow substitution, but not elimination, reactions to occur rapidly under conditions relevant to carbon capture. Furthermore, the sudden adsorption of CO₂ in this compound is thus associated with a transition from a dynamic surface state to a well-ordered extended surface structure. Accordingly, the reaction with CO₂ can be considered to be thermodynamically non-spontaneous at low pressures because of the large decrease in entropy associated with this transition. Indeed, the molar entropy of gas-phase CO₂ was found to be the primary determinant of the step pressure for phase-change adsorbents. As shown in Extended Data Fig. 4, step pressures for all five phase-change metal organic frameworks were linearly correlated with the gas-phase entropy of CO₂ as a function of temperature.

Understanding and manipulating the isotherm steps

The mechanism of CO₂ adsorption suggests that variation of the metal–amine bond strength should provide a method of manipulating the

isotherm step position. The series of isostructural compounds $\text{mmen-M}_2(\text{dobpdc})$ ($M = \text{Mg, Mn, Fe, Co, Ni, Zn}$) were therefore synthesized, and the CO_2 adsorption isotherms for each were measured at 25, 40, 50 and 75 °C (Fig. 5). With the exception of the Ni compound, which showed normal Langmuir-type adsorption behaviour (Fig. 5e), all of the materials showed sharp isotherm steps that shifted to higher pressure with increasing temperature. Analysis of the isotherm steps at 25 °C yielded Hill coefficients²³ of 10.6, 5.6, 7.5, 11.5 and 6.0 for $M = \text{Mg, Mn, Fe, Co}$ and Zn , respectively, reflecting the cooperative nature of the CO_2 adsorption mechanism. Simulated isotherms generated from grand-canonical Monte Carlo simulations using a simple lattice model captured the experimentally observed isotherm step only when all mmen groups reacted with CO_2 and aligned down the crystallographic c axis (see Extended Data Fig. 5).

For a given temperature, the step position varies in the order $\text{Mg} < \text{Mn} < \text{Fe} < \text{Zn} < \text{Co}$, in good agreement with the published series for octahedral metal complex stabilities²⁴. The lack of a step for the Ni compound, even at very high pressures (Extended Data Fig. 6) is attributable to the exceptional stability of the Ni– mmen bond, which prevents carbamate insertion from taking place under the conditions surveyed. Geometry optimizations performed with periodic DFT calculations²⁵ using various functionals were in good agreement with experimentally calculated values, and the trend in the calculated adsorption energies was directly correlated with the calculated metal–amine bond length (Extended Data Fig. 5). Thus, we predict that similar variations in tuning step position will be possible for the $\text{M}_2(\text{dobpdc})$ series by altering the sterics of the amine bound to the metal, as well as the spacer between the two amine groups. Hence, depending on the concentration of CO_2

present in a gas mixture, an adsorbent can be rationally designed to match the optimum process conditions depicted in Fig. 1.

Although stepped adsorption isotherms²⁶ have been observed previously in solid adsorbents, the origin of the step reported here is unique and distinct from all previously reported mechanisms. First, in contrast to most metal–organic frameworks showing such behaviour, the isotherm steps reported here are not attributable to pore-opening, gate-opening or pore-closing processes^{27–29}. For $\text{mmen-Mn}_2(\text{dobpdc})$, only a $\sim 1\%$ decrease in the unit cell volume was observed on CO_2 adsorption, and from Fig. 3a, b it is apparent that the entire pore surface was accessible to CO_2 throughout the adsorption isotherm. A gate-opening mechanism attributable to the rearrangement of flexible hydrogen-bonding functional groups, which function by preventing CO_2 diffusion into the pores at low partial pressures³⁰, cannot explain the presence of distinct adsorption steps when the material is slowly cooled from high to low temperatures under isobaric adsorption conditions (see below). Second, in contrast to adsorbed-layer phase transitions on highly homogeneous surfaces, the adsorbed phase reported here was stable at temperatures well above the critical temperature of CO_2 (ref. 31). Third, the phase transition was a metal cation-dependent, solid-to-solid transformation, in contrast to liquid-to-liquid or liquid-to-solid phase change reactions typically reported for amine– CO_2 systems^{32–34}. Last, under conditions relevant to CO_2 capture, desorption hysteresis was minimal, because the sharp steps occurred over a narrow pressure regime and the adsorption and desorption onset points were at about the same temperature and pressure (see Extended Data Fig. 7).

Several features unique to the $\text{mmen-M}_2(\text{dobpdc})$ series permitted phase transitions of this type to be observed. First, for solid ammonium carbamate chains to form, the metal–amine coordinate bond must be capable of rearrangement. Thus, only amines tethered to the solid surface through coordinate bonds rather than covalent bonds can undergo the mechanism reported here. Second, a homogeneous surface with appropriately positioned adsorption sites, which is dictated by the location of open metal sites within the pores of the metal–organic framework, is necessary. Thus, a very limited number of metal–organic framework materials would be able to mimic the adsorption behaviour reported here, and it is likely that no amine-functionalized mesoporous silica sorbent could be engineered precisely enough to meet these requirements. Notably, in contrast to the pore expanded derivatives of $\text{M}_2(\text{dobdc})$ reported here, amine functionalization of the parent $\text{Mg}_2(\text{dobdc})$ compound was not reported to result in stepped adsorption isotherms³⁵.

Low-energy carbon capture applications

Effective adsorbents for carbon capture must possess large working capacities for processes occurring at temperatures above 40 °C and at CO_2 partial pressures near 0.15 bar for coal flue gas or near 0.05 bar for a natural gas flue stream. On this basis, the location of the isotherm steps for the Mg and Mn compounds makes them better suited for this application than the Fe, Co or Zn compounds, which are better suited for separations from gas mixtures with higher CO_2 concentrations. To assess the utility of these phase-change adsorbents for capturing CO_2 in a pure temperature swing adsorption process, adsorption isobars were collected under dynamic gas flow. Samples of $\text{mmen-Mg}_2(\text{dobpdc})$ and $\text{mmen-Mn}_2(\text{dobpdc})$ were activated, saturated with 100% CO_2 and then cooled isobarically to room temperature under three different CO_2 -containing gas mixtures: 100%, 15% and 5%. The resulting isobars, shown in Fig. 6a, b, reveal how small changes in temperature induced large changes in the quantity of CO_2 adsorbed. As shown in Fig. 6c, d, phase-change adsorbents showed very large working capacities when used in temperature swing adsorption processes. For $\text{mmen-Mg}_2(\text{dobpdc})$ to give a working capacity in excess of 13 wt%, the material must simply swing between 100 and 150 °C. Similarly, the working capacity of $\text{mmen-Mn}_2(\text{dobpdc})$ was in excess of 10 wt% when cycled between 70 and 120 °C. In particular, to simulate a pure temperature swing adsorption process accurately, 15% CO_2 in N_2 was flowed over the samples during the cooling phase, whereas 100% CO_2 was used during heating phases.

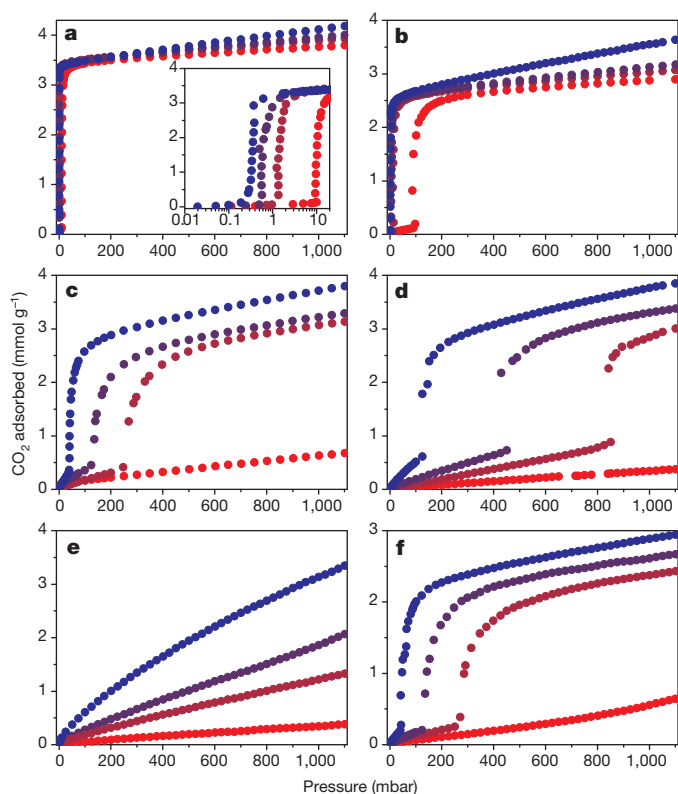


Figure 5 | CO_2 adsorption isotherms. Carbon dioxide adsorption isotherms at 25 °C (blue), 40 °C (blue-violet), 50 °C (red-violet) and 75 °C (red) for $\text{mmen-Mg}_2(\text{dobpdc})$ (a), $\text{mmen-Mn}_2(\text{dobpdc})$ (b), $\text{mmen-Fe}_2(\text{dobpdc})$ (c), $\text{mmen-Co}_2(\text{dobpdc})$ (d), $\text{mmen-Ni}_2(\text{dobpdc})$ (e) and $\text{mmen-Zn}_2(\text{dobpdc})$ (f). Despite the use of aliphatic amine groups as the CO_2 reactive species, the metal–organic framework has an essential role in determining isotherm shape, owing to the importance of metal–ligand reorganization reactions in the mechanism.

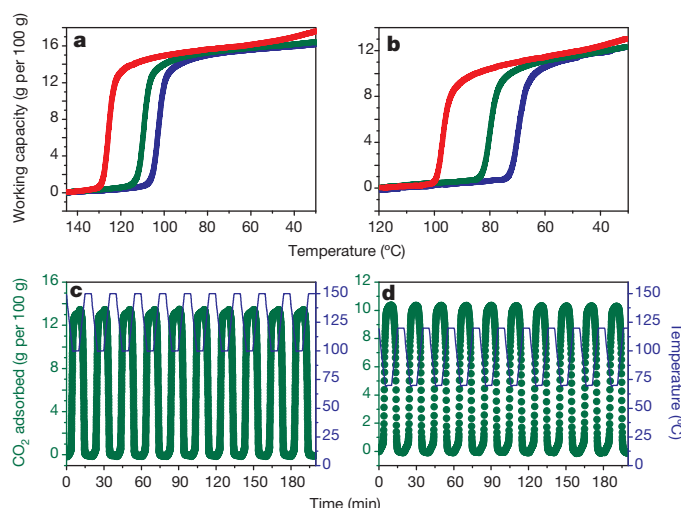


Figure 6 | Isobaric CO₂ adsorption and cycling experiments. **a, b,** Variable-temperature adsorption isobars of 100% (red), 15% (green) and 5% (blue) CO₂ (in N₂) for mmen-Mg₂(dobpdc) (**a**) and mmen-Mn₂(dobpdc) (**b**), showing that under dynamic conditions the sharp transition region allows phase-change adsorbents to achieve very large working capacities under a wide range of adsorption conditions. For each material, the phase-transition temperature is dependent on the pressure of CO₂ in the gas mixture, with higher phase-transition temperatures being observed at higher CO₂ partial pressures. **c, d,** Cycling data for a pure temperature swing process involving adsorption from a simulated coal flue gas (15% CO₂ in N₂) at 100 °C in mmen-Mg₂(dobpdc) (**c**) and 70 °C in mmen-Mn₂(dobpdc) (**d**), followed by desorption at 150 and 120 °C, respectively, using a flow of 100% CO₂. Respective working capacities of 13% and 10% are attained, with no loss in capacity over the course of ten cycles.

In contrast to experiments that use a purge gas to assist CO₂ desorption, no inert gases were used to regenerate the samples.

Estimations based on differential scanning calorimetry, thermogravimetric analysis and isosteric heat determinations indicate that mmen-Mg₂(dobpdc) and mmen-Mn₂(dobpdc) can achieve regeneration energies of about 2.2–2.3 MJ per kg of CO₂ captured. This value is appreciably lower than the regeneration energies attainable using monoethanolamine (3.5 MJ kg^{−1}) or even state-of-the-art amines, such as piperazine and KS-1 (2.6 MJ kg^{−1})^{5,36}. In contrast to aqueous amine absorbents that use heat exchangers to save sensible energy costs, the greater working capacities and smaller temperature swings of phase-change adsorbents allow more economical processes to be developed for a high-enthalpy adsorbent without the use of a heat exchanger. Because phase-change adsorbents saturate with CO₂ at their transition point, it is not necessary for adsorption to occur at the lowest possible temperature. Whereas we previously showed that mmen-Mg₂(dobpdc) can operate effectively under standard flue gas adsorption conditions (40 °C)¹³, Fig. 6 shows that phase-change adsorbents operated more efficiently at higher adsorption temperatures than at lower temperatures. Because classical adsorbents must operate at the lowest possible adsorption temperature to maximize working capacity, only phase-change adsorbents can enable high-temperature adsorption processes to be considered.

Adsorbing CO₂ at elevated temperatures affords several additional process benefits besides directly decreasing sorbent regeneration energy. In particular, overcoming the competitive adsorption of water vapour, which is present in flue gas at high concentrations, presents a serious challenge for solid adsorbents. Amine-based solid adsorbents fare better than those using a purely physical adsorption mechanism, because they are known to retain their affinity for CO₂ under humid conditions³⁷, as also shown here for mmen-Mg₂(dobpdc) (see Extended Data Fig. 7 for additional dynamic gas adsorption experiments). However, even for systems where the amine reactivity with CO₂ is unaffected by the presence of water, the physical adsorption of water on non-amine

binding sites increases the overall regeneration energy of the material³⁸. As shown in Extended Data Fig. 6b, mmen-Mg₂(dobpdc) adsorbed nearly 90% less water at 100 °C than at 40 °C. Thus, the energy penalty associated with desorbing co-adsorbed water can be substantially decreased by performing CO₂ adsorption at a high temperature, obviating the need for strict flue gas dehumidification. No changes to the CO₂ adsorption isotherm were apparent after exposure to water at 40 or 100 °C, indicating the stability of the mmen-Mg₂(dobpdc) in the presence of water vapour even at high temperatures.

The high effective operating temperatures of mmen-Mg₂(dobpdc) and mmen-Mn₂(dobpdc) offer opportunities for cost savings beyond just decreases in the regeneration energy. Because of the exothermic nature of all adsorption processes, the incorporation of labour and material intensive coolant pipes into an adsorbent bed (a component of the considerable infrastructure cost for carbon capture) is necessary to maintain isothermal adsorption conditions. The rate of heat transfer from a sorbent bed to the coolant pipes, which contain surface temperature water at ~25 °C, is primarily dependent on the heat transfer coefficient of the sorbent, the total contact area between the sorbent and the coolant pipes, and the temperature differential between the sorbent and the coolant³⁹. The physical size of adsorption units is dictated, to a great extent, by the need to provide sufficient contact area between the coolant and sorbent for effective heat removal. For processes that are limited by heat transfer rather than mass transfer, which is likely for many CO₂ capture processes using solid adsorbents, the use of high temperatures will maximize the temperature differential between the coolant and the sorbent, substantially reducing the overall bed size by reducing the size of the necessary contact area. By increasing the coolant–sorbent temperature differential from about 15 °C to nearly 75 °C, adsorption bed size could potentially be reduced fivefold. In turn, smaller adsorbent beds would reduce the pressure drop across the adsorbent, reduce the size and cost of the required capital equipment, and allow as little as one-fifth as much adsorbent to be used. By decreasing these other system costs, new classes of adsorbents have the ability to reduce the cost of carbon capture substantially beyond simply decreasing the sorbent regeneration energy.

A functional model for Rubisco

The reactivity trends of the M₂(dobpdc) series may help to clarify the evolutionary conservation of Mg²⁺ within the active site of most photosynthetic enzymes. Biological fixation of atmospheric CO₂ is effected primarily by the ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) enzyme. Striking structural similarities exist between mmen-Mg₂(dobpdc) and the enzymatic pocket of Rubisco, which in its active form also contains an octahedral Mg²⁺ ion ligated by five oxygen donor ligands and a reactive aliphatic amine ligand that adsorbs gas-phase CO₂ to form an O-bound carbamate ligand (Extended Data Fig. 1f)^{40,41}. Although other divalent metal ions can be incorporated into either structure, in each case the presence of Mg²⁺ greatly enhances the reactivity for CO₂ fixation at very low CO₂ concentrations⁴². Although further study is necessary, the trends that we observed suggest that the inclusion of Mg²⁺ within the active site of Rubisco may be necessary to endow the lysine residue that forms the enzymatically competent carbamate-ligated metal with sufficient reactivity at low partial pressures of CO₂.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 September 2014; accepted 6 February 2015.

Published online 11 March 2015.

1. International Energy Agency. *CO₂ Emissions from Fuel Combustion: Highlights* <<http://www.iea.org/publications/freepublications/publication/CO2EmissionsFromFuelCombustionHighlights2013.pdf>> (IEA, 2013).
2. IPCC. in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. et al.) 11–14 (Cambridge Univ. Press, 2013).

3. Orr, J. C. *et al.* Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. *Nature* **437**, 681–686 (2005).
4. Haszeldine, R. S. Carbon capture and storage: how green can black be? *Science* **325**, 1647–1652 (2009).
5. Boot-Handford, M. E. *et al.* Carbon capture and storage update. *Energy Environ. Sci.* **7**, 130–189 (2014).
6. Choi, S., Drese, J. H. & Jones, C. W. Adsorbent materials for carbon dioxide capture from large anthropogenic point sources. *ChemSusChem* **2**, 796–854 (2009).
7. Lin, L. C. *et al.* *In silico* screening of carbon-capture materials. *Nature Mater.* **11**, 633–641 (2012).
8. Sumida, K. *et al.* Carbon dioxide capture in metal-organic frameworks. *Chem. Rev.* **112**, 724–781 (2012).
9. Zhou, H. C., Long, J. R. & Yaghi, O. M. Introduction to metal-organic frameworks. *Chem. Rev.* **112**, 673–674 (2012).
10. Furukawa, H., Cordova, K. E., O’Keeffe, M. & Yaghi, O. M. The chemistry and applications of metal-organic frameworks. *Science* **341**, 123044 (2013).
11. Demessence, A., D’Alessandro, D. M., Foo, M. L. & Long, J. R. Strong CO₂ binding in water stable triazolate-bridged metal-organic framework functionalized with ethylenediamine. *J. Am. Chem. Soc.* **131**, 8784–8786 (2009).
12. McDonald, T. M., D’Alessandro, D. M., Krishna, R. & Long, J. R. Enhanced carbon dioxide capture upon incorporation of *N,N'*-dimethylethylenediamine in the metal-organic framework CuBTTri. *Chem. Sci.* **2**, 2022–2028 (2011).
13. McDonald, T. M. *et al.* Capture of carbon dioxide from air and flue gas in the alkylamine-appended metal-organic framework mmen-Mg₂(dobpdc). *J. Am. Chem. Soc.* **134**, 7056–7065 (2012).
14. Hong, C. S. *et al.* Diamine-functionalized metal-organic framework: exceptionally high CO₂ capacities from ambient air and flue gas, ultrafast CO₂ uptake rate, and adsorption mechanism. *Energy Environ. Sci.* **7**, 744–751 (2014).
15. Rosi, N. L. *et al.* Rod packings and metal-organic frameworks constructed from rod-shaped secondary building units. *J. Am. Chem. Soc.* **127**, 1504–1518 (2005).
16. Dietzel, P. D. C., Panella, B., Hirscher, M., Blom, R. & Fjellvåg, H. Hydrogen adsorption in a nickel based coordination polymer with open metal sites in the cylindrical cavities of the desolvated framework. *Chem. Commun.* 959–961 (2006).
17. Caskey, S. R., Wong-Foy, A. G. & Matzger, A. J. Dramatic tuning of carbon dioxide uptake via metal substitution in a coordination polymer with cylindrical pores. *J. Am. Chem. Soc.* **130**, 10870–10871 (2008).
18. Mason, J. A. *et al.* Evaluating metal-organic frameworks for post-combustion carbon dioxide capture via temperature swing adsorption. *Energy Environ. Sci.* **4**, 3030–3040 (2011).
19. Coelho, A. A. Whole-profile structure solution from powder diffraction powder using simulated annealing. *J. Appl. Crystallogr.* **33**, 899–908 (2000).
20. Planas, N. *et al.* The mechanism of carbon dioxide adsorption in an alkylamine-functionalized metal-organic framework. *J. Am. Chem. Soc.* **135**, 7402–7405 (2013).
21. Tiritiris, I. & Kantelehn, W. Orthoamide und Iminiumsalze, LXX [1]. Zur Fixierung von Kohlendioxid mit organischen Basen (Teil 1): Reaktionen von Diaminen mit Kohlendioxid. *Z. Naturforsch.* **66b**, 164–176 (2011).
22. Drisdell, W. S. *et al.* Probing adsorption interactions in metal-organic frameworks using X-ray spectroscopy. *J. Am. Chem. Soc.* **135**, 18183–18190 (2013).
23. Weiss, J. N. The Hill equation revisited: uses and misuses. *FASEB J.* **11**, 835–841 (1997).
24. Irving, H. & Williams, R. J. P. The stability of transition-metal complexes. *J. Chem. Soc.* **637**, 3192–3210 (1953).
25. Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
26. Walton, K. S. *et al.* Understanding inflections and steps in carbon dioxide adsorption isotherms in metal-organic frameworks. *J. Am. Chem. Soc.* **130**, 406–407 (2008).
27. Horike, S., Shimomura, S. & Kitagawa, S. Soft porous crystals. *Nature Chem.* **1**, 695–704 (2009).
28. Serre, C. *et al.* Very large breathing effect in the first nanoporous chromium(III)-based solids: MIL-53 or Cr^{III}(OH)(O₂C–C₆H₄–CO₂)₂·(HO₂C–C₆H₄–CO₂H)_x·(H₂O)_y. *J. Am. Chem. Soc.* **124**, 13519–13526 (2002).
29. Henke, S. *et al.* Multiple phase-transitions upon selective CO₂ adsorption in an alkyl ether functionalized metal-organic framework—an *in situ* X-ray diffraction study. *CrystEngComm* **13**, 6399–6404 (2011).
30. Seo, J., Matsuda, R., Sakamoto, H., Bonneau, C. & Kitagawa, S. A pillared-layer coordination polymer with a rotatable pillar acting as a molecular gate for guest molecules. *J. Am. Chem. Soc.* **131**, 12792–12800 (2009).
31. Thomy, A. & Duval, X. Stepwise isotherms and phase transitions in physisorbed films. *Surf. Sci.* **299–300**, 415–425 (1994).
32. Jessop, P. G., Mercer, S. M. & Heldebrand, D. J. CO₂-triggered switchable solvents, surfactants, and other materials. *Energy Environ. Sci.* **5**, 7240–7253 (2012).
33. Liebenthal, U. *et al.* Overall process analysis and optimisation for CO₂ capture from coal fired power plants based on phase change solvents forming two liquid phases. *Energy Procedia* **37**, 1844–1854 (2013).
34. Ma’mun, S. & Kim, I. Selection and characterization of phase-change solvent for carbon dioxide capture: precipitating system. *Energy Procedia* **37**, 331–339 (2013).
35. Choi, S., Watanabe, T., Bae, T.-H., Sholl, D. S. & Jones, C. W. Modification of the Mg/DOBDC MOF with amines to enhance CO₂ adsorption from ultradilute gases. *J. Phys. Chem. Lett.* **3**, 1136–1141 (2012).
36. Rochelle, G. *et al.* Aqueous piperazine as the new standard for CO₂ capture technology. *Chem. Eng. J.* **171**, 725–733 (2011).
37. Sayari, A. & Blemabkhout, Y. Stabilization of amine-containing CO₂ adsorbents: dramatic effect of water vapor. *J. Am. Chem. Soc.* **132**, 6312–6314 (2010).
38. Nugent, P. *et al.* Porous materials with optimal adsorption thermodynamics and kinetics for CO₂ separation. *Nature* **495**, 80–84 (2013).
39. van Lare, C. E. J. *Mass Transfer in Gas Fluidized Beds: Scaling, Modeling, and Particle Size Influence* 141–142. PhD thesis, Tech. Univ. Eindhoven (1991) <<http://alexandria.tue.nl/repository/books/348157.pdf>>.
40. Lorimer, G. The carboxylation and oxygenation of ribulose 1,5-bisphosphate: the primary events in photosynthesis and photorespiration. *Annu. Rev. Plant Physiol.* **32**, 349–382 (1981).
41. Taylor, T. C. & Andersson, I. Structural transitions during activation and ligand binding in hexadecameric Rubisco inferred from the crystal structure of activated unliganded spinach enzyme. *Nature Struct. Biol.* **3**, 95–101 (1996).
42. Assche, F. & Clijsters, H. Effects of metals on enzyme activity in plants. *Plant Cell Environ.* **13**, 195–206 (1990).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. S. Bhowan and A. H. Berger of EPRI, H. Krutka, C. M. Brown and K. S. Suslick for discussions, and L. Ribaud and the 11-BM staff at the Advanced Photon Source at Argonne National Laboratory for assisting with powder X-ray diffraction experiments. The work presented here pertaining to the synthesis and gas adsorption properties of metal-organic frameworks was funded by the Advanced Research Projects Agency–Energy (ARPA-E), US Department of Energy (DOE), under award numbers DE-AR0000103 and DE-AR0000402. Funding pertaining to the characterization of materials by spectroscopy and X-ray diffraction and the computational work performed by W.S.D., B.V., R.P., S.K.S., K.L., J.B.N., B.S. and J.B.K. was provided by the Center for Gas Separations Relevant to Clean Energy Technologies, an Energy Frontier Research Center funded by the DOE, Office of Science, Office of Basic Energy Sciences under award DE-SC0001015. Experiments performed in Turin were supported by grant MIUR-PRIN 2010-2011. Work at SIMAP was performed using computer resources from GENCI (CINES grant 2014-c2015097211). The computational work performed by S.O.O., A.L.D., N.P. and L.G. was supported through the Nanoporous Materials Genome Center of the DOE, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences, under award number DE-FG02-12ER16362. This research used resources of the Advanced Photon Source, a DOE Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. Portions of this work (use of beamline 6.3.2 at the Advanced Light Source; a user project at The Molecular Foundry, facilitated by T.P., L.F.W. and D.P., and use of its computer cluster vulcan, managed by the High Performance Computing Services Group; use of the National Energy Research Scientific Computing Center) were performed at Lawrence Berkeley National Laboratory, which is supported by the Director, Office of Science, Office of Basic Energy Sciences, of the DOE under contract no. DE-AC02-05CH11231. For fellowship support, we further thank the National Science Foundation (J.A.M.), Gerald K. Branch and Arkema (E.D.B.) and the Research Council of Norway (grant 230534 to S.K.S.).

Author Contributions T.M.McD. and J.R.L. formulated the project. T.M.McD., E.D.B. and D.G. synthesized the compounds. T.M.McD. collected and analysed the gas adsorption data. J.A.M. collected and analysed the X-ray diffraction data. X.K. collected the NMR data. X.K. and J.A.R. analysed the NMR data. T.M.McD., A.D., V.C., F.G. and S.B. collected and analysed the infrared data. W.S.D. and J.B.K. collected X-ray absorption spectroscopy data and performed analysis, with assistance from R.P., T.P., L.F.W. and D.P. S.O.O., B.V., A.L.D., R.P., S.K.S., N.P. and K.L. performed the computations and analysed the results. J.B.N., B.S. and L.G. helped with the computational analyses. T.M.McD., J.A.M. and J.R.L. wrote the paper, and all authors contributed to revising the paper.

Author Information Metrical data for the solid-state structures of mmen-Mn₂(dobpdc) without and with adsorbed CO₂ are available free of charge from the Cambridge Crystallographic Data Centre under reference numbers CCDC 994497 and 994498, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.R.L. (jrlong@berkeley.edu).

METHODS

General synthesis and characterization methods. All reagents and solvents were obtained from commercial sources at reagent-grade purity or higher. The compound $\text{H}_4(\text{dobpdc})$ was synthesized as reported previously¹³. No statistical methods were used to predetermine sample size.

Laboratory powder X-ray diffraction patterns were collected on a Bruker AXS D8 Advance diffractometer equipped with $\text{Cu K}\alpha$ radiation ($\lambda = 1.5418 \text{ \AA}$), a Göbel mirror, a Lynxeye linear position-sensitive detector, and mounting the following optics: fixed divergence slit (0.6 mm), receiving slit (3 mm) and secondary-beam Soller slits (2.5°). The generator was set at 40 kV and 40 mA. Owing to the oxygen sensitivity of $\text{Fe}_2(\text{dobpdc})$ and $\text{mmen-Fe}_2(\text{dobpdc})$, X-ray diffraction patterns were collected in sealed glass capillaries placed on the powder stage. Thermogravimetric analysis was carried out at a ramp rate of $2^\circ \text{C min}^{-1}$ in a nitrogen flow with a TA Instruments Q5000. Elemental analyses for C, H and N were performed at the Microanalytical Laboratory of the University of California, Berkeley.

Synthesis of $\text{Mg}_2(\text{dobpdc})$. To a 20 ml glass scintillation vial, $\text{H}_4(\text{dobpdc})$ (27.4 mg, 0.10 mmol), $\text{Mg}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ (64.0 mg, 0.25 mmol) and 10 ml of mixed solvent (55:45 methanol/dimethylformamide (DMF)) were added. The vial was sealed with a polytetrafluoroethylene (PTFE)-lined cap and placed in a well plate 2 cm deep on a 393 K hot plate. After 12 h a white powder formed on the bottom and walls of the vial. The reaction mixture was then decanted and the remaining powder was soaked three times in DMF and then three times in methanol. The solid was then collected by filtration and fully desolvated by heating under dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 523 K for 24 h to afford 23.3 mg (0.073 mmol), 73% of $\text{Mg}_2(\text{dobpdc})$. Combustion elemental analysis calculated (Anal. Calcd) for $\text{C}_{14}\text{H}_6\text{O}_6\text{Mg}_2$: C, 52.74; H, 1.90. Found: C, 52.47; H, 1.64.

Synthesis of $\text{Mn}_2(\text{dobpdc})$. To a 20 ml glass scintillation vial, $\text{H}_4(\text{dobpdc})$ (27.4 mg, 0.10 mmol), $\text{MnCl}_2 \cdot 4\text{H}_2\text{O}$ (49.5 mg, 0.25 mmol), and 10 ml of mixed solvent (1:1 ethanol/DMF) were added. The vial was sealed with a PTFE-lined cap and placed in a well plate 2 cm deep on a 393 K hot plate. After 12 h a yellow powder formed on the bottom and walls of the vial. The reaction mixture was then decanted and the remaining powder was soaked three times in DMF and then three times in methanol. The solid was then collected by filtration and fully desolvated by heating under dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 523 K for 24 h to afford 33.8 mg (0.0889 mmol), 89% of $\text{Mn}_2(\text{dobpdc})$. Anal. Calcd for $\text{C}_{14}\text{H}_6\text{O}_6\text{Mn}_2$: C, 44.24; H, 1.59. Found: C, 44.32; H, 1.23.

Synthesis of $\text{Fe}_2(\text{dobpdc})$. Anhydrous FeCl_2 (2.85 g, 22.4 mmol), $\text{H}_4(\text{dobpdc})$ (1.85 g, 6.75 mmol), anhydrous DMF (400 ml) and anhydrous methanol (50 ml) were added to a 500 ml Schlenk flask under an argon atmosphere. The reaction mixture was heated to 393 K and stirred for 24 h to afford a dark yellow-green precipitate. The solvent was then removed by cannula transfer and replaced with fresh anhydrous DMF. The reaction mixture was soaked three times in DMF and then three times in methanol. The solid was then fully desolvated by heating under dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 523 K for 24 h to afford 2.395 g (6.28 mmol), 93% of $\text{Fe}_2(\text{dobpdc})$. Anal. Calcd for $\text{C}_{14}\text{H}_6\text{O}_6\text{Fe}_2$: C, 44.03; H, 1.58. Found: C, 43.72; H, 1.48.

Synthesis of $\text{Co}_2(\text{dobpdc})$. To a 20 ml glass scintillation vial, $\text{H}_4(\text{dobpdc})$ (41.1 mg, 0.15 mmol), $\text{Co}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ (109 mg, 0.375 mmol) and 15 ml of mixed solvent (1:1 water/DMF/ethanol) were added. The vial was sealed with a PTFE-lined cap and placed in a well plate 2 cm deep on a 393 K hot plate. After 36 h a pink powder formed on the bottom of the vial. The reaction mixture was then decanted and the remaining powder was soaked three times in DMF and then three times in methanol. The solid was then collected by filtration and fully desolvated by heating under dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 523 K for 24 h to afford 54.1 mg (0.139 mmol), 93% of $\text{Co}_2(\text{dobpdc})$. Anal. Calcd for $\text{C}_{14}\text{H}_6\text{O}_6\text{Co}_2$: C, 43.33; H, 1.56. Found: C, 42.92; H, 1.38.

Synthesis of $\text{Zn}_2(\text{dobpdc})$. To a 20 ml glass scintillation vial, $\text{H}_4(\text{dobpdc})$ (27.4 mg, 0.10 mmol), $\text{ZnBr}_2 \cdot 2\text{H}_2\text{O}$ (83.5 mg, 0.32 mmol) and 10 ml of mixed solvent (1:1 ethanol/DMF) were added. The vial was sealed with a PTFE-lined cap and placed in a well plate 2 cm deep on a 393 K hot plate. After 12 h a pale yellow powder formed on the bottom and walls of the vial. The reaction mixture was then decanted and the remaining powder was soaked three times in DMF and then three times in methanol. The solid was then collected by filtration and fully desolvated by heating under dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 523 K for 24 h to afford 21.4 mg (0.0534 mmol), 53% of $\text{Zn}_2(\text{dobpdc})$. Anal. Calcd for $\text{C}_{14}\text{H}_6\text{O}_6\text{Zn}_2$: C, 41.94; H, 1.51. Found: C, 41.26; H, 1.57.

Synthesis of $\text{Ni}_2(\text{dobpdc})$. To a 20 ml glass scintillation vial, $\text{H}_4(\text{dobpdc})$ (41.1 mg, 0.15 mmol), $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$ (109 mg, 0.375 mmol) and 15 ml of mixed solvent (1:1 water/DMF/ethanol) were added. The vial was sealed with a PTFE-lined cap and placed in a well plate 2 cm deep on a 393 K hot plate. After 36 h, a green powder formed on the bottom of the vial. The reaction mixture was then decanted and the remaining powder was soaked three times in DMF and then three times in methanol. The solid was then collected by filtration and fully desolvated by heating under

dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 523 K for 24 h to afford 39.3 mg (0.101 mmol), 68% of $\text{Ni}_2(\text{dobpdc})$. Anal. Calcd for $\text{C}_{14}\text{H}_6\text{O}_6\text{Ni}_2$: C, 43.39; H, 1.56. Found: C, 43.09; H, 1.24.

General synthesis of $\text{mmen-M}_2(\text{dobpdc})$. In a plastic glovebag with positive N_2 pressure, 10 ml of 10% mmen solution in hexanes was added to $\sim 100 \text{ mg}$ of activated $\text{M}_2(\text{dobpdc})$ in a glass Micromeritics adsorption tube. The tube was sealed with a rubber septum and left to sit undisturbed for 4 h in the glovebag. In the glovebag, the sample was collected by vacuum filtration and rinsed with five 10 ml portions of dry hexanes. The hexanes-solvated sample was desolvated under dynamic vacuum ($<10 \text{ }\mu\text{bar}$) at 348 K (for Zn) or 373 K (for Mg, Mn, Fe, Co and Ni) for 4 h.

Elemental analysis of $\text{mmen-M}_2(\text{dobpdc})$ series. Mg: Anal. Calcd for $\text{C}_{22}\text{H}_{30}\text{N}_4\text{O}_6\text{Mg}_2$: C, 53.37; H, 6.11; N, 11.32. Found: C, 52.39; H, 5.52; N, 10.36. Mn: $\text{C}_{22}\text{H}_{30}\text{N}_4\text{O}_6\text{Mn}_2$: C, 47.47; H, 5.43; N, 10.07. Found: C, 47.26; H, 5.20; N, 10.24. Fe: Anal. Calcd for $\text{C}_{22}\text{H}_{30}\text{N}_4\text{O}_6\text{Fe}_2$: C, 47.34; H, 5.42; N, 10.04. Found: C, 47.3; H, 4.95; N, 9.71. Co: Anal. Calcd for $\text{C}_{22}\text{H}_{30}\text{N}_4\text{O}_6\text{Co}_2$: C, 46.82; H, 5.36; N, 9.93. Found: C, 46.27; H, 4.94; N, 9.61. Zn: Anal. Calcd for $\text{C}_{22}\text{H}_{26}\text{N}_4\text{O}_6\text{Zn}_2$: C, 45.78; H, 5.24; N, 9.71. Found: C, 45.46; H, 4.75; N, 9.78. Ni: Anal. Calcd for $\text{C}_{22}\text{H}_{30}\text{N}_4\text{O}_6\text{Ni}_2$: C, 46.86; H, 5.36; N, 9.94. Found: C, 46.68; H, 5.33; N, 10.28.

Low-pressure gas adsorption measurements. For all low-pressure (0–1.1 bar) gas adsorption measurements, 60–130 mg of adsorbent was transferred to a pre-weighed glass sample tube under an atmosphere of nitrogen gas and capped with a Transeal. Samples were then transferred manually to a Micromeritics ASAP 2020 gas adsorption analyser and heated to the activation temperatures previously specified, under vacuum. The sample was considered activated when the outgas rate was less than $2 \text{ }\mu\text{bar min}^{-1}$. The evacuated tube containing the activated sample was then transferred to a balance and weighed to determine the mass of the desolvated sample. The tube was then placed manually on the analysis port of the aforementioned instrument, where the outgas rate was once again confirmed to be less than $2 \text{ }\mu\text{bar min}^{-1}$. Isothermal conditions were maintained at 77 K with liquid N_2 , at 25, 40, 50 and 75°C with a Julabo F32 water circulator, and at 100°C with a heated sand bath controlled by a programmable temperature controller.

High-pressure gas adsorption measurements. The high-pressure CO_2 adsorption isotherm for $\text{mmen-Ni}_2(\text{dobpdc})$ was measured on a HPVA-II-100 from Particulate Systems, a Micromeritics company. Here, 0.27 g of activated $\text{mmen-Ni}_2(\text{dobpdc})$ was loaded into a tared 2 ml stainless steel sample holder inside a glovebox under a N_2 atmosphere. Before the sample holder was connected to the variable compression ratio fittings of the high-pressure assembly inside the glovebox, the sample holder was weighed to determine the sample mass.

Before $\text{mmen-Ni}_2(\text{dobpdc})$ was measured, 25°C CO_2 background measurements were performed on a sample holder containing nonporous glass beads that occupied a similar volume as a typical sample. A small negative background was observed at higher pressures, which may have been due to errors in volume calibrations, temperature calibrations and/or the equation of state used to perform the nonideality corrections. Nevertheless, the background CO_2 adsorption was consistent across several measurements and was well described by fitting to a polynomial equation. This polynomial was then used to perform a background subtraction on the raw high-pressure CO_2 data for $\text{mmen-Ni}_2(\text{dobpdc})$.

In situ infrared spectroscopy. A powdered sample of $\text{mmen-Mg}_2(\text{dobpdc})$ ($\sim 15 \text{ mg}$) was pelletized, shaped in a self-supported wafer and placed inside a commercial Fourier-transform infrared reactor cell (2000-A multimode; AABSPEC), which allowed infrared spectra to be recorded under flow conditions at a wide range of temperatures. Before CO_2 was flowed across the sample, the sample was first activated at 150°C for 15 min while flowing 30 ml min^{-1} of pure nitrogen (heating ramp rate of $1^\circ \text{C min}^{-1}$). Next, the gas flow was switched to 5% CO_2 in N_2 at a flow rate of 30 ml min^{-1} , and the system was cooled from 150°C to 30°C at a rate of $1^\circ \text{C min}^{-1}$. Before cooling, the sample was conditioned for 15 min at 150°C with the 5% CO_2 in N_2 gas mixture. After cooling to 30°C , the sample was heated again to 150°C at a rate of $5^\circ \text{C min}^{-1}$ under N_2 , to check the reversibility of the process. Spectra were collected every 5°C with a resolution of 2 cm^{-1} (number of scans equal to 32) on a Perkin-Elmer System 2000 infrared spectrophotometer equipped with a HgCdTe detector. For Extended Data Fig. 2c, curve-fitting analysis was performed with the Levenberg–Marquardt method by using the OPUS software (Bruker Optik); 100% Gauss functions were used.

For Extended Data Fig. 2b, the attenuated total reflectance (ATR) accessory of a Perkin-Elmer Spectrum 400 was enclosed within a plastic glovebag filled with positive gas pressure. Spectra of activated adsorbent were collected first under a N_2 atmosphere. After 32 scans at a resolution of 4 cm^{-1} , the glovebag was allowed to fill with CO_2 for 5 min and infrared spectra of the sample were collected again under an atmosphere of CO_2 . For Extended Data Fig. 2e, a gas mixture of 20 mbar H_2O , 150 mbar CO_2 and 600 mbar N_2 was dosed onto an activated sample of $\text{mmen-Mg}_2(\text{dobpdc})$ and was left to equilibrate for 6 h before being placed on the ATR stage of a glovebag-encased Spectrum 400 instrument.

Solid-state NMR spectroscopy. Solid-state NMR experiments were performed on a 7.05 T magnet with a Tecmag Discovery spectrometer, using a Doty 5 mm triple-resonance magic angle spinning probe. The frequency of ^{15}N was 30.4 MHz. ^{15}N chemical shifts were referenced to ^{15}N -labelled glycine at 33 p.p.m. relative to liquid ammonia. The experiments were performed at ambient temperature. Magic angle spinning was used to collect high-resolution NMR spectra with a spinning rate ranging from 5 to 6 kHz. The 90° pulses for ^1H , ^{13}C and ^{15}N were 3.8, 4.7 and 7 μs , respectively. Ramped cross-polarization with variable contact times from 1 to 10 ms was used to generate ^{15}N signals. Recycle delays were set to be fivefold longer than the T_1 of protons. Two-pulse phase-modulated ^1H decoupling of 100 kHz was applied during ^{15}N signal acquisition.

Isobaric CO_2 adsorption methods. Isobaric gravimetric adsorption experiments were collected using a TA Instruments Q5000 analyser using premixed gas cylinders (Praxair). A flow rate of 25 ml min^{-1} was employed for all gases. Before each experiment, the samples (~ 5 mg) were desolvated by heating under N_2 for 1 h. To simulate temperature swing adsorption processes accurately, samples were activated and reactivated with only 100% CO_2 . Sample masses were normalized to be 0% under a 100% CO_2 atmosphere ~ 10 s before the onset of each cooling cycle. Thus, the base mass corresponds to the weight of the metal-organic framework and any CO_2 adsorbed under the activation conditions: 150 $^\circ\text{C}$ for mmen- $\text{Mg}_2(\text{dobpdc})$ and 120 $^\circ\text{C}$ for mmen- $\text{Mn}_2(\text{dobpdc})$. Masses were uncorrected for buoyancy effects.

For Fig. 6a, b, ramp rates of 1 $^\circ\text{C min}^{-1}$ were used; samples were reactivated at the appropriate regeneration temperature under 100% CO_2 for 15 min between gases. Switching from 100% CO_2 to a lower concentration of CO_2 occurred at the onset of cooling, ensuring that the lower-concentration gas would not simulate a purge gas for regeneration.

For Fig. 6c, d, ramp rates of 10 $^\circ\text{C min}^{-1}$ were used. Samples were heated between the adsorption and regeneration conditions under 100% CO_2 ; the regeneration time was 5 min. Samples were cooled under low-concentration CO_2 ; the adsorption time was 5 min. Total cycle time was ~ 20 min.

High-resolution powder X-ray diffraction. Samples of fully activated mmen- $\text{Mn}_2(\text{dobpdc})$ microcrystalline powders (~ 10 mg) were loaded into 1.0 mm boron-rich glass capillaries inside a glovebox under an N_2 atmosphere. The capillaries were attached to a gas cell, which was connected to the analysis port of a Micromeritics ASAP-2020 gas adsorption instrument. The capillaries were fully evacuated at room temperature for 30 min, dosed with 5 mbar of He [mmen- $\text{Mn}_2(\text{dobpdc})$], 5 mbar of CO_2 (100K- CO_2 -mmen- $\text{Mn}_2(\text{dobpdc})$) or 100 mbar of CO_2 (295K- CO_2 -mmen- $\text{Mn}_2(\text{dobpdc})$) and then equilibrated at room temperature for 15 min, 8 h or 4 h, respectively. After equilibration, the capillaries were flame-sealed and placed inside a Kapton tube that was sealed on both ends with epoxy.

High-resolution synchrotron X-ray powder diffraction data were subsequently collected at beamline 11-BM at the Advanced Photon Source (APS), Argonne National Laboratory, with an average wavelength of ~ 0.4137 Å. Diffraction patterns were collected at 100 K, 100 K and 295 K for mmen- $\text{Mn}_2(\text{dobpdc})$, 100K- CO_2 -mmen- $\text{Mn}_2(\text{dobpdc})$ and 295K- CO_2 -mmen- $\text{Mn}_2(\text{dobpdc})$, respectively. Discrete detectors covering an angular range from -6 to $16^\circ 2\theta$ were scanned over a $34^\circ 2\theta$ range, with data points collected every $0.001^\circ 2\theta$ and a scan speed of 0.01°s^{-1} . Owing to the large number of collected data points, all diffraction patterns were rebinned to a step size of $0.005^\circ 2\theta$ before structure solution and Rietveld refinement. Additionally, all diffraction patterns showed a high-intensity peak at $\sim 1.25^\circ$ that was $\sim 500\%$ more intense than any other diffraction peak. Because this high d -spacing peak does not contribute important structural information and was heavily biasing all structure solution attempts, all data analysis was performed with a minimum 2θ of 2° .

A standard peak search, followed by indexing through the Single Value Decomposition approach⁴³, as implemented in TOPAS-Academic⁴⁴, allowed the determination of approximate unit-cell parameters. Tentatively, the space groups for both mmen- $\text{Mn}_2(\text{dobpdc})$ and CO_2 -mmen- $\text{Mn}_2(\text{dobpdc})$ were assigned as $P3_221$ because the framework was expected to be isostructural to $\text{Zn}_2(\text{dobpdc})$, which was previously characterized by single crystal X-ray diffraction¹³. Precise unit-cell dimensions were determined by structureless Le Bail refinements (Supplementary Table 1). Here, the background was modelled by a polynomial function of the Chebyshev type, and anisotropic peak broadening was described using parameters appropriate for a hexagonal crystal system⁴⁵. Successful structure solution and Rietveld refinement confirmed that $P3_221$ was indeed the correct space group for all compounds.

Breakthrough adsorption measurements. Into a glass U-tube with an interior diameter of 4 mm, 203 mg of activated mmen- $\text{Mg}_2(\text{dobpdc})$ powder was added. The column was placed inside a furnace with a programmable temperature controller. Gas flow rates of 10 ml min^{-1} were used. Column effluent was analysed using a Hy-Energy/Setaram RGAPro-2500 with continuous sampling capability. Argon (99.999%; Praxair) was used as a purge gas for sample activation at 100 $^\circ\text{C}$ for 30 min. For measurements at 25 $^\circ\text{C}$, $\sim 1.5\%$ H_2O was added to the gas stream by bubbling a premixed gas cylinder of 15% CO_2 , 4% O_2 and balance N_2 (Praxair) through a glass bubbler containing distilled water.

NEXAFS measurements. *In situ* NEXAFS measurements were performed with a gas cell instrument previously described in detail elsewhere^{22,46}. Transmission-absorption samples were prepared by suspending and sonicating ~ 10 mg of mmen- $\text{Mg}_2(\text{dobpdc})$ powder in ~ 1 ml of hexanes, and drop casting onto SiC membranes (300 nm thickness). The metal-organic framework-coated membranes were then loaded into a glass evacuation chamber that was slowly evacuated to 50 mTorr. The chamber was then heated to 100 $^\circ\text{C}$ for 4 h to remove all guest species from the pores. After heating, the chamber was cooled to ambient temperature and refilled with dry nitrogen gas. Samples were transferred from the evacuation chamber to the gas cell in a dry nitrogen glovebox to prevent exposure to any unwanted species. N K-edge spectra of mmen- $\text{Mg}_2(\text{dobpdc})$ were collected at beamline 6.3.2 (10^{11} photons s^{-1}) at the Advanced Light Source at Lawrence Berkeley National Laboratory, under vacuum and increasing pressures of CO_2 gas. The experimental procedure was the same as in earlier studies. The full pressure series of N K-edge spectra are shown in Extended Data Fig. 3.

AIMD simulations. The equilibrium structure of mmen- $\text{Mg}_2(\text{dobpdc})$ is obtained by performing constant-pressure (NPT) *ab initio* molecular dynamics simulations at room temperature. The system is equilibrated using a Parrinello–Rahman barostat and a Langevin thermostat with a time step of 0.5 fs (refs 47, 48). A plane-wave basis set that is truncated at 400 eV is used to represent the electronic wavefunctions within the projector-augmented wave (PAW) approximation^{49,50}, as realized in the Vienna *Ab initio* Simulation Package (VASP)^{51,52}. vdW-DF2 functional is used to approximate the long-range dispersion forces⁵³, which is potentially important to capture the weak intermolecular interactions between the metal-organic framework and the adsorbed molecules. Here, the vdW-DF2 correlation is computed with the use of a $1 \times 1 \times 3$ k -point grid to ensure that the electron density is converged for a $22 \text{ Å} \times 21 \text{ Å} \times 7 \text{ Å}$ supercell. On equilibration, eight time-separated (that is, uncorrelated) snapshots are taken from the molecular dynamics trajectory to establish an average spectrum from NEXAFS simulations.

XAS simulations. DFT calculations used for X-ray absorption spectroscopy (XAS) simulations employ the PBE GGA functional²⁵. Plane-wave pseudopotential calculations using ultrasoft pseudopotentials⁵⁴ and a kinetic energy cutoff for electronic wavefunctions (density) of 25 (200) Rydberg (Ry) were performed using the PWSCF code within the Quantum-ESPRESSO package⁵⁵. The core-excited Kohn–Sham eigenspectrum was generated using the eXcited electron and Core Hole (XCH) approach. On the basis of a numerically converged self-consistent charge density, we generated the unoccupied states for our XAS calculations non-self-consistently, sufficiently sampling the first Brillouin zone with a $2 \times 2 \times 2$ uniform k -point grid, employing an efficient implementation of the Shirley interpolation scheme⁵⁶ generalized to handle ultrasoft pseudopotentials⁵⁷. Matrix elements were evaluated within the PAW frozen-core approximation⁵⁸. Core-excited ultrasoft pseudopotentials and corresponding atomic orbitals were generated with the Vanderbilt code⁵⁴. Each computed transition was convoluted with a 0.2 eV Gaussian function to produce continuous spectra.

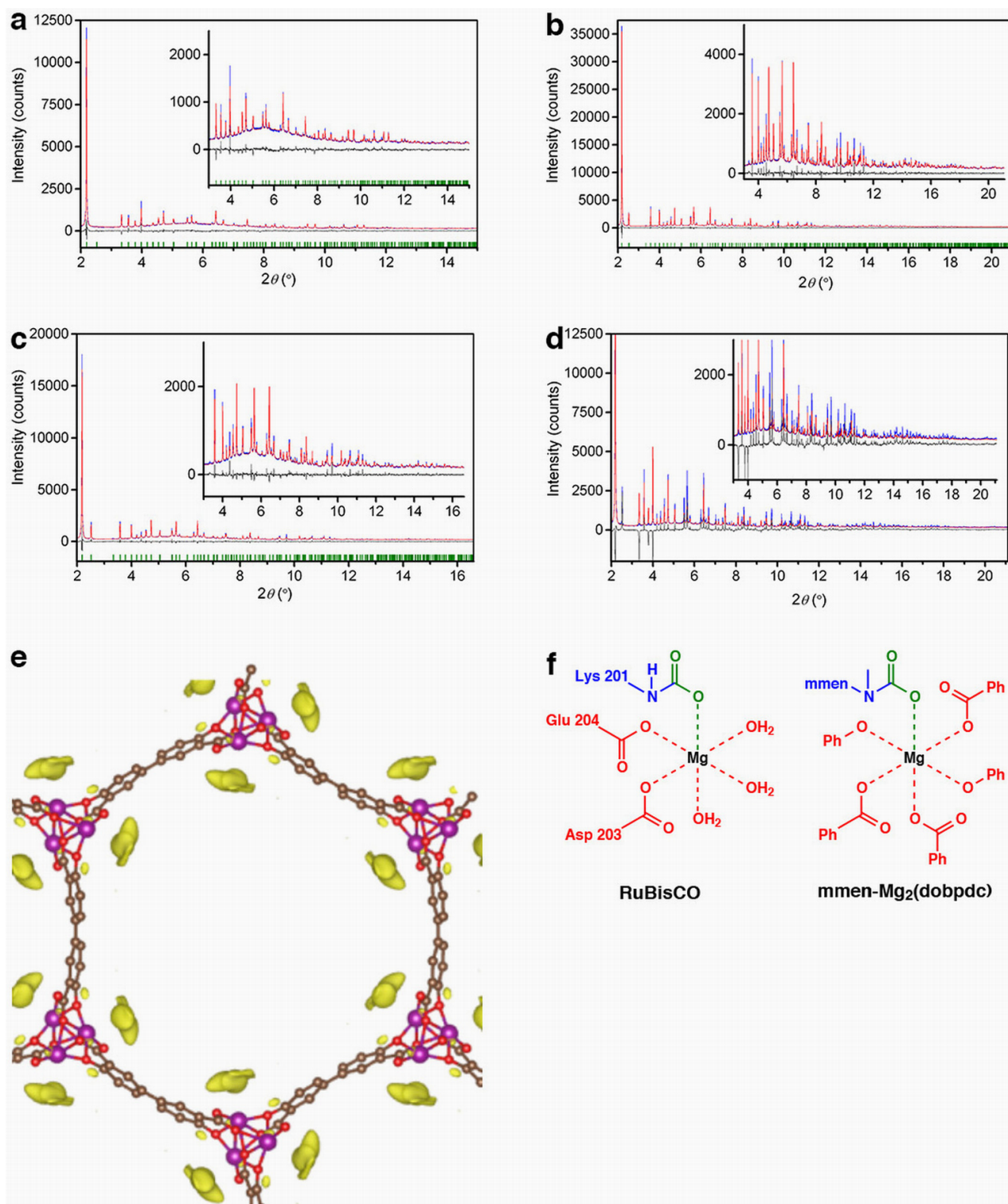
Periodic DFT calculations. Periodic DFT calculations in this work were performed with the VASP 5.3.3 package^{51,59}. The energetics of CO_2 capture by the alkylamine moieties chemisorbed in the channels of $\text{M}_2(\text{dobpdc})$ ($\text{M} = \text{Mg}, \text{Mn}, \text{Fe}, \text{Co}, \text{Ni}, \text{Zn}$) were computed with the PBE and M06L functionals^{25,60}. The electron–ion interactions in these calculations were described with the PAW method⁴⁹ with an energy cutoff of 550 eV. Atomic positions and lattice parameters were optimized until the forces on all atoms were smaller than 0.02 eV Å^{-1} at the Γ -point. On-site Hubbard U corrections were employed for metal d electrons for $\text{M}_2(\text{dobpdc})$ ($\text{M} = \text{Mn}, \text{Fe}, \text{Co}, \text{Ni}$)⁶¹. The U values are determined to reproduce oxidation energies in the respective metal oxides⁶².

Modelled adsorption isotherms. To study the differences in adsorption behaviour between the pair and chain models, we used a lattice model to predict the adsorption isotherms. The lattice models are illustrated in Extended Data Fig. 5. The energy at each lattice point is determined by the state of the lattice point and the state of the surrounding lattice points. For both the pair and the chain model the interaction energies, yellow and red, respectively, are taken directly from the DFT calculations. The end of a chain (shown in blue) is 80% of the chain model, whereas a single adsorbed CO_2 (shown in green) is 10% of the energy of a chain. Similarly, interactions between two rows (cross-channel) are set to the pair energy (or 75% of a chain). An amine without CO_2 was not given an energy contribution. The effect of having different metals was taken into account by varying the energy of the chain in accordance with the DFT energies. To compute the isotherms we performed standard grand-canonical Monte Carlo simulations. To compare the lattice model chemical potential directly with the chemical potential of CO_2 we used a shift of the pressure, which was fitted to the steps of the isotherms at the highest and lowest temperatures.

43. Coelho, A. A. Indexing of powder diffraction patterns by iterative use of singular value decomposition. *J. Appl. Crystallogr.* **36**, 86–95 (2003).

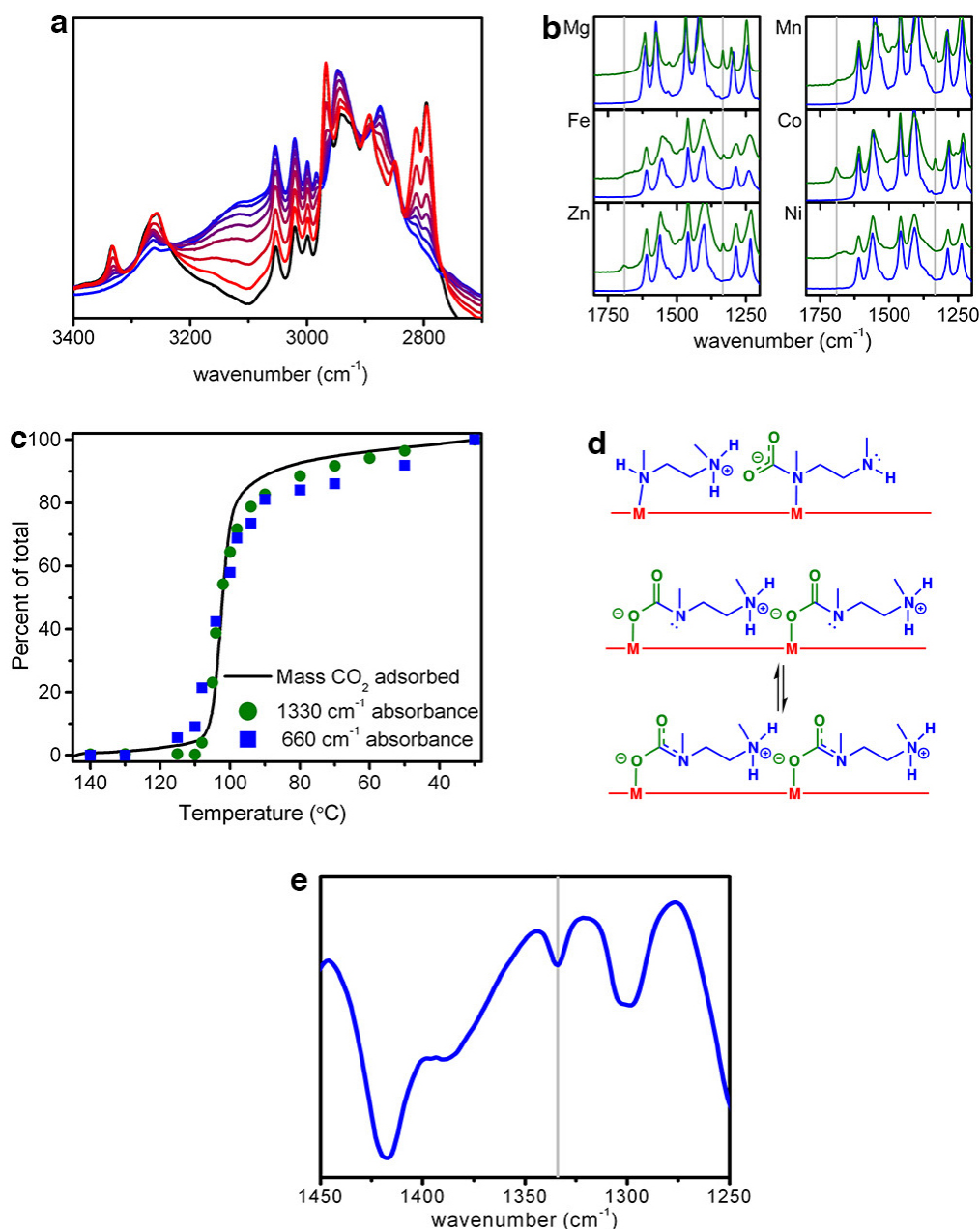
44. Coelho, A. A. TOPAS-Academic, Version 4.1 (Coelho Software, 2007).

45. Stephens, P. W. Phenomenological model of anisotropic peak broadening in power diffraction. *J. Appl. Crystallogr.* **32**, 281–289 (1999).
46. Drisdell, W. S. & Kortright, J. B. Gas cell for *in situ* soft X-ray transmission-absorption spectroscopy of materials. *Rev. Sci. Instrum.* **85**, 074103 (2014).
47. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids* (Clarendon Press, 1987).
48. Parrinello, M. & Rahman, A. Crystal structure and pair potentials: a molecular-dynamics study. *Phys. Rev. Lett.* **45**, 1196–1199 (1980).
49. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17978 (1994).
50. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
51. Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
52. Kresse, G. & Hafner, J. Ab initio molecular-dynamics for liquid-metals. *Phys. Rev. B* **47**, 558–561 (1993).
53. Lee, K., Murray, E. D., Kong, L., Lundqvist, B. I. & Langreth, D. C. Higher-accuracy van der Waals density functional. *Phys. Rev. B* **82**, 081101 (2010).
54. Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892–7895 (1990).
55. Giannozzi, P. *et al.* QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).
56. Shirley, E. L. Optimal basis sets for detailed Brillouin-zone integrations. *Phys. Rev. B* **54**, 16464–16469 (1996).
57. Prendergast, D. & Louie, S. G. Bloch-state-based interpolation: an efficient generalization of the Shirley approach to interpolating electronic structure. *Phys. Rev. B* **80**, 235126 (2009).
58. Taillefumier, M., Cabaret, D., Flank, A.-M. & Mauri, F. X-ray absorption near-edge structure calculations with the pseudopotentials: application to the K edge in diamond and α -quartz. *Phys. Rev. B* **66**, 195107 (2002).
59. Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
60. Zhao, Y. & Truhlar, D. G. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.* **125**, 194101 (2006).
61. Liechtenstein, A. I., Anisimov, V. I. & Zaanen, J. Density-functional theory and strong interactions: orbital ordering in Mott–Hubbard insulators. *Phys. Rev. B* **52**, R5467–R5470 (1995).
62. Wang, L., Maxisch, T. & Ceder, G. Oxidation energies of transition metal oxides within the GGA+U framework. *Phys. Rev. B* **73**, 195107 (2006).



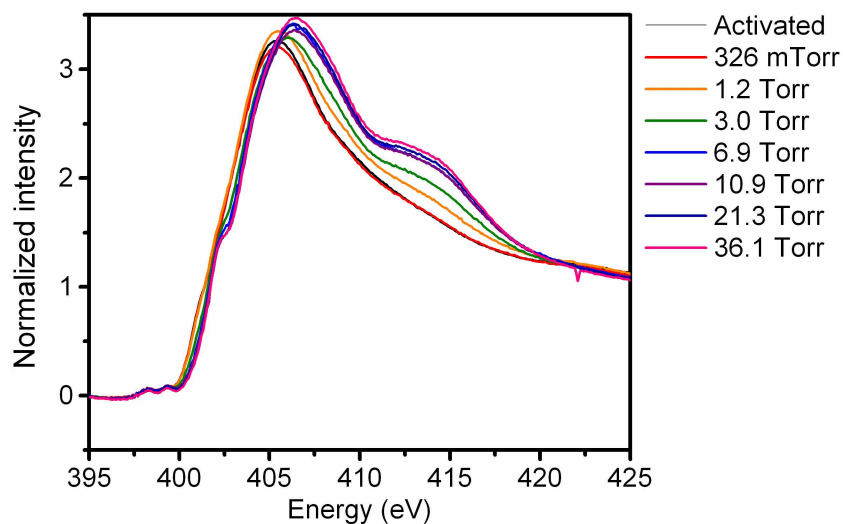
Extended Data Figure 1 | Powder X-ray diffraction. **a–c**, Rietveld refinement of $\text{mmen-Mn}_2(\text{dobpdc})$ at 100 K (**a**), $\text{CO}_2\text{-mmen-Mn}_2(\text{dobpdc})$ at 100 K (**b**) and $\text{CO}_2\text{-mmen-Mn}_2(\text{dobpdc})$ at 295 K (**c**). The blue and red lines represent the experimental and calculated diffraction patterns, respectively; the grey line represents the difference between experimental and calculated patterns; the green tick marks represent the calculated Bragg peak positions. **d**, Plot of the diffraction data for $\text{CO}_2\text{-mmen-Mn}_2(\text{dobpdc})$ at 100 K (blue), where the calculated pattern (red) is based on the $\text{mmen-Mn}_2(\text{dobpdc})$ structural model. The grey line represents the difference between the

experimental and calculated patterns. Note that the intensity differences indicate a structural transition when CO_2 is adsorbed. **e**, Fourier difference map for $\text{mmen-Mn}_2(\text{dobpdc})$ at 100 K. Purple, brown and red spheres represent Mn, C and O atoms, respectively; yellow blobs represent excess electron density that is not accounted for in the $\text{Mn}_2(\text{dobpdc})$ structural model and that is due to the mmen bound to each Mn^{2+} site. **f**, The coordination environment around Mg^{2+} in the active form of the Rubisco enzyme is structurally similar to the coordination environment around the metal cations of $\text{mmen-Mg}_2(\text{dobpdc})$ after adsorption of CO_2 .



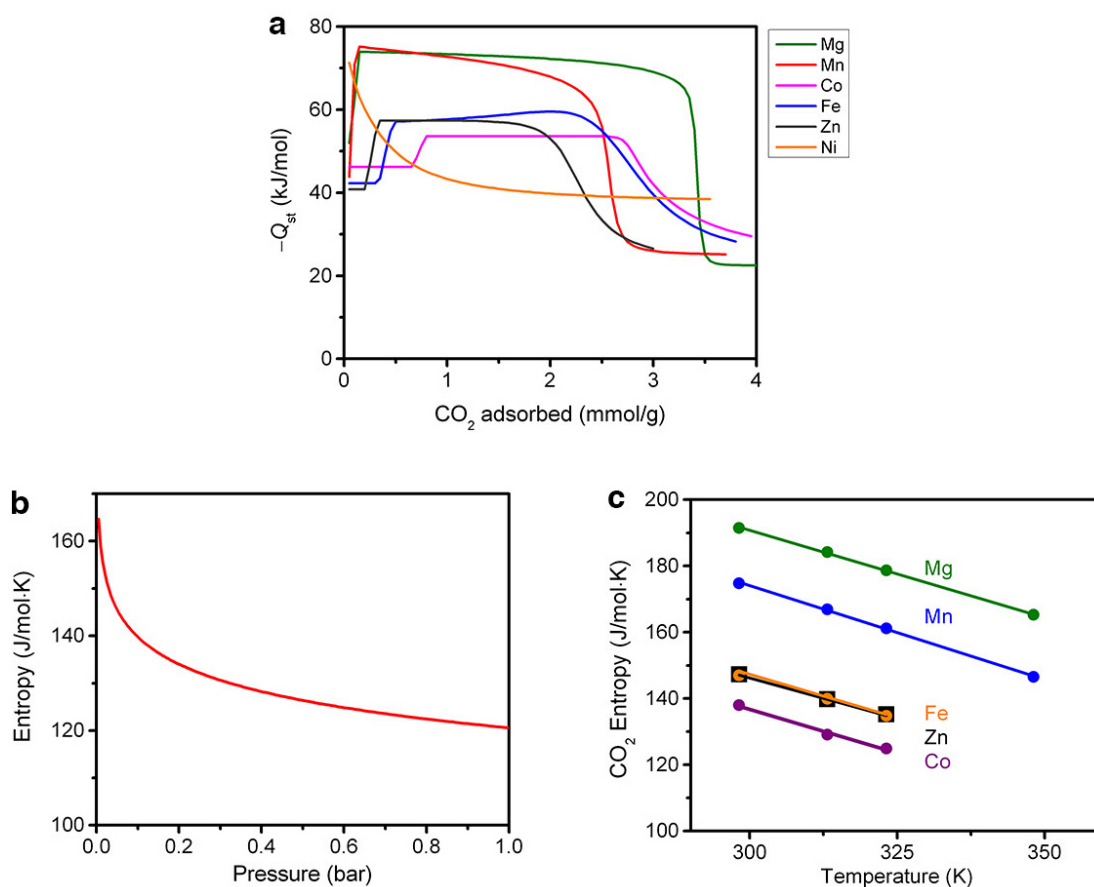
Extended Data Figure 2 | Infrared spectroscopy. **a**, On cooling mmen- $\text{Mg}_2(\text{dobpdc})$ from $150\text{ }^{\circ}\text{C}$ (red) to $30\text{ }^{\circ}\text{C}$ (blue) under 5% CO_2 , changes to the aliphatic C–H vibrations of mmen are apparent on adsorption of CO_2 . Furthermore, ammonium formation from neutral secondary amines is indicated by the appearance of a new, broad feature centred on $3,000\text{ cm}^{-1}$. **b**, Room temperature, *in situ* infrared spectroscopy measurements of mmen- $\text{M}_2(\text{dobpdc})$ ($\text{M} = \text{Mg}, \text{Mn}, \text{Fe}, \text{Co}, \text{Zn}, \text{Ni}$) under N_2 (blue) and CO_2 (green) atmospheres. Grey lines mark diagnostic carbamate bands at $1,690$ and $1,334\text{ cm}^{-1}$. **c**, On cooling under flowing 5% CO_2 in N_2 from 150 to $30\text{ }^{\circ}\text{C}$, the normalized mass increase of mmen- $\text{Mg}_2(\text{dobpdc})$ measured by

thermogravimetric analysis (black line) can be compared with the normalized integrated area of the infrared active bands at $1,330$ and 660 cm^{-1} . The bands at $1,330$ and 660 cm^{-1} can be assigned to $\nu(\text{C–N})$ and $[\beta(\text{OCO}) + \beta(\text{NCO})]$ modes that are characteristic of the highly ordered ammonium carbamate chains. **d**, In contrast to carbamate that is coordinated to a metal site through a nitrogen atom (top), changes in electron resonance configurations give rise to a feature at $1,334\text{ cm}^{-1}$ characteristic of CO_2 insertion into the metal–nitrogen bond (bottom). **e**, Infrared spectroscopy clearly indicates that CO_2 adsorption via metal–amine insertion occurs even in the presence of H_2O owing to the presence of the aforementioned band at $1,334\text{ cm}^{-1}$.



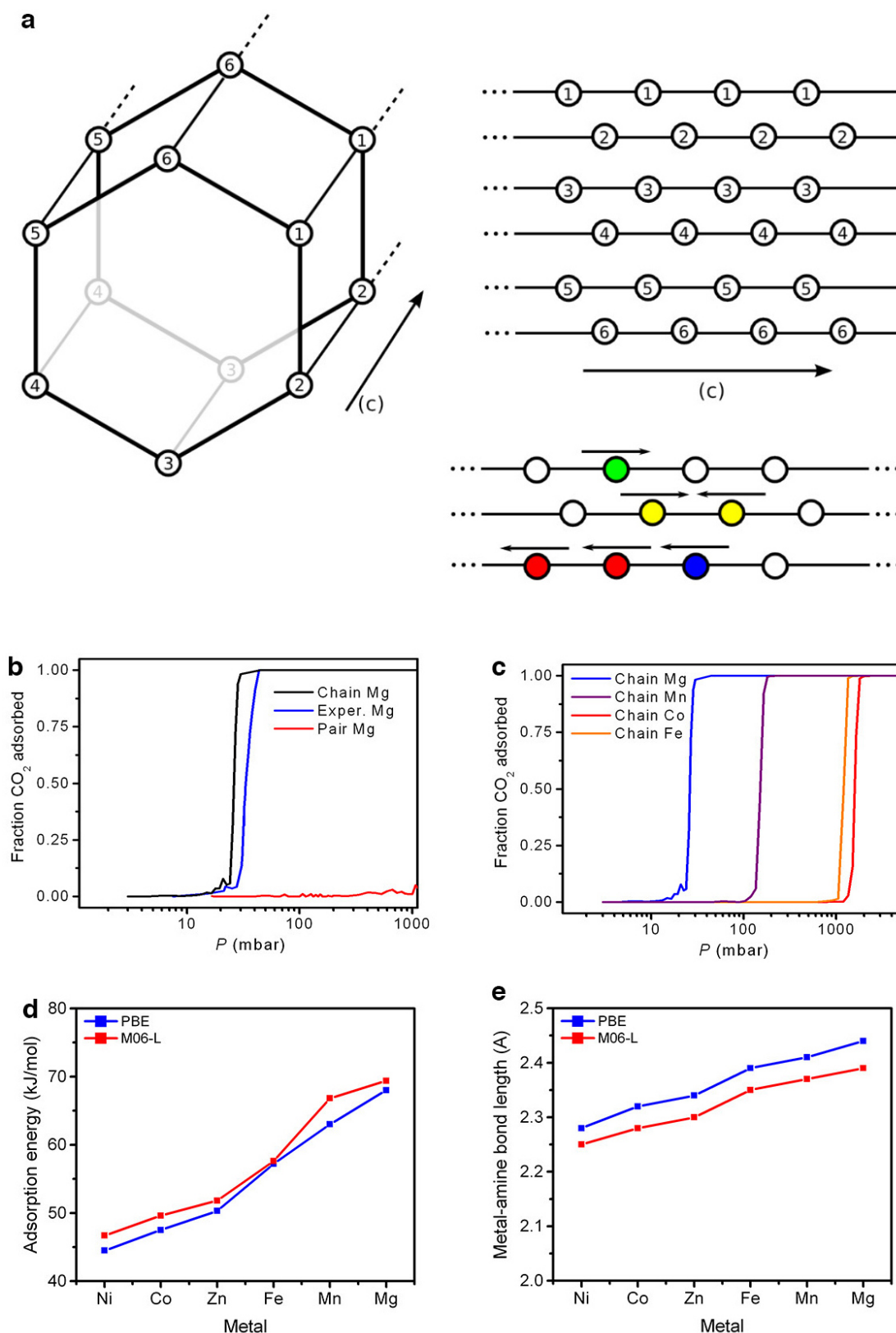
Extended Data Figure 3 | X-ray absorption spectroscopy. Experimental N K-edge NEXAFS spectra of mmen-Mg₂(dobpdc) in vacuum and under increasing pressures of CO₂ gas. The broad feature between 411 and 419 eV, a

signature of N–C bond formation, appears before the pre-edge peak at 402.3 eV, which is characteristic of CO₂ insertion.



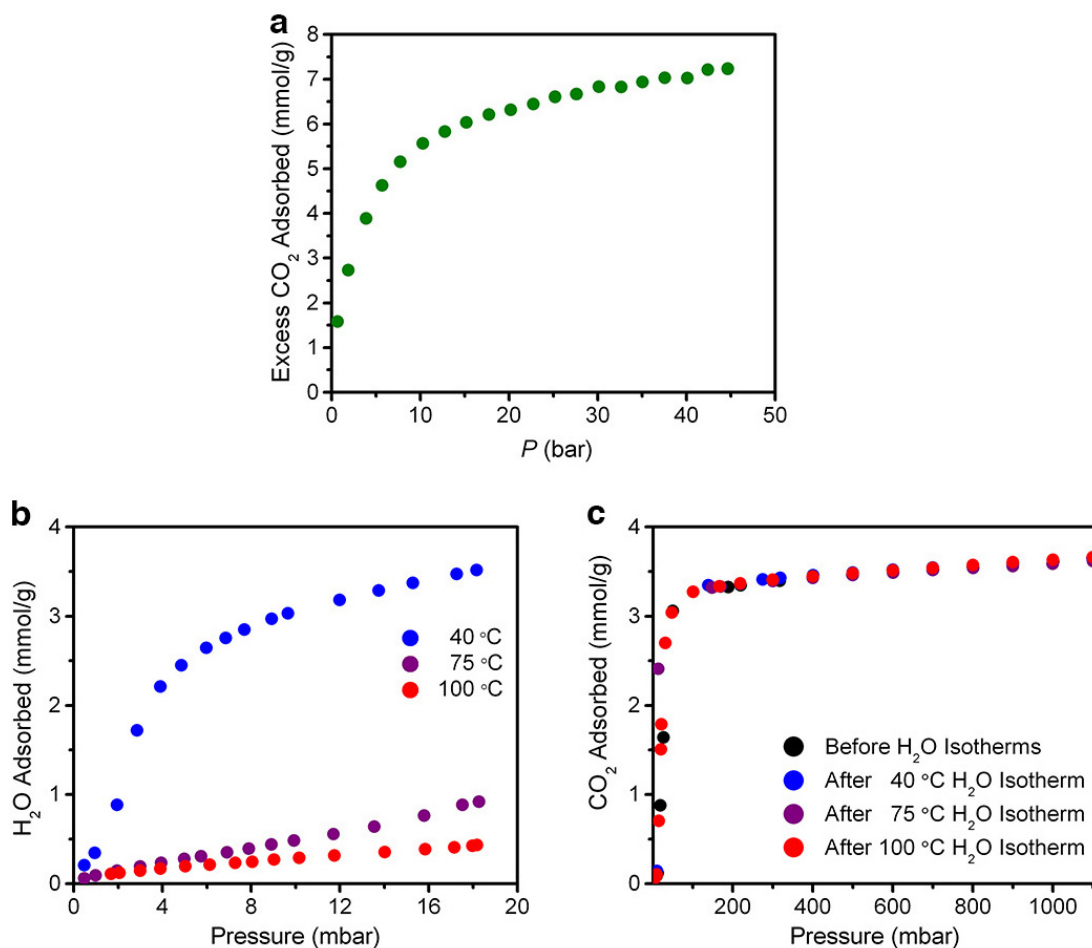
Extended Data Figure 4 | Thermodynamics of CO₂ adsorption. **a**, Isosteric heat of CO₂ adsorption plots for the mmen-M₂(dobpdc) series. **b**, Plot of the entropy of gaseous CO₂ against pressure at 298 K. **c**, A linear correlation was found to exist for each mmen-M₂(dobpdc) material between the step pressure at any temperature and the gas-phase entropy of CO₂, for

mmen-Mg₂(dobpdc) (green, $R^2 = 0.99946$), mmen-Mn₂(dobpdc) (blue, $R^2 = 0.99918$), mmen-Fe₂(dobpdc) (orange, $R^2 = 0.99934$), mmen-Co₂(dobpdc) (purple, $R^2 = 0.99244$) and mmen-Zn₂(dobpdc) (black, $R^2 = 0.99932$).



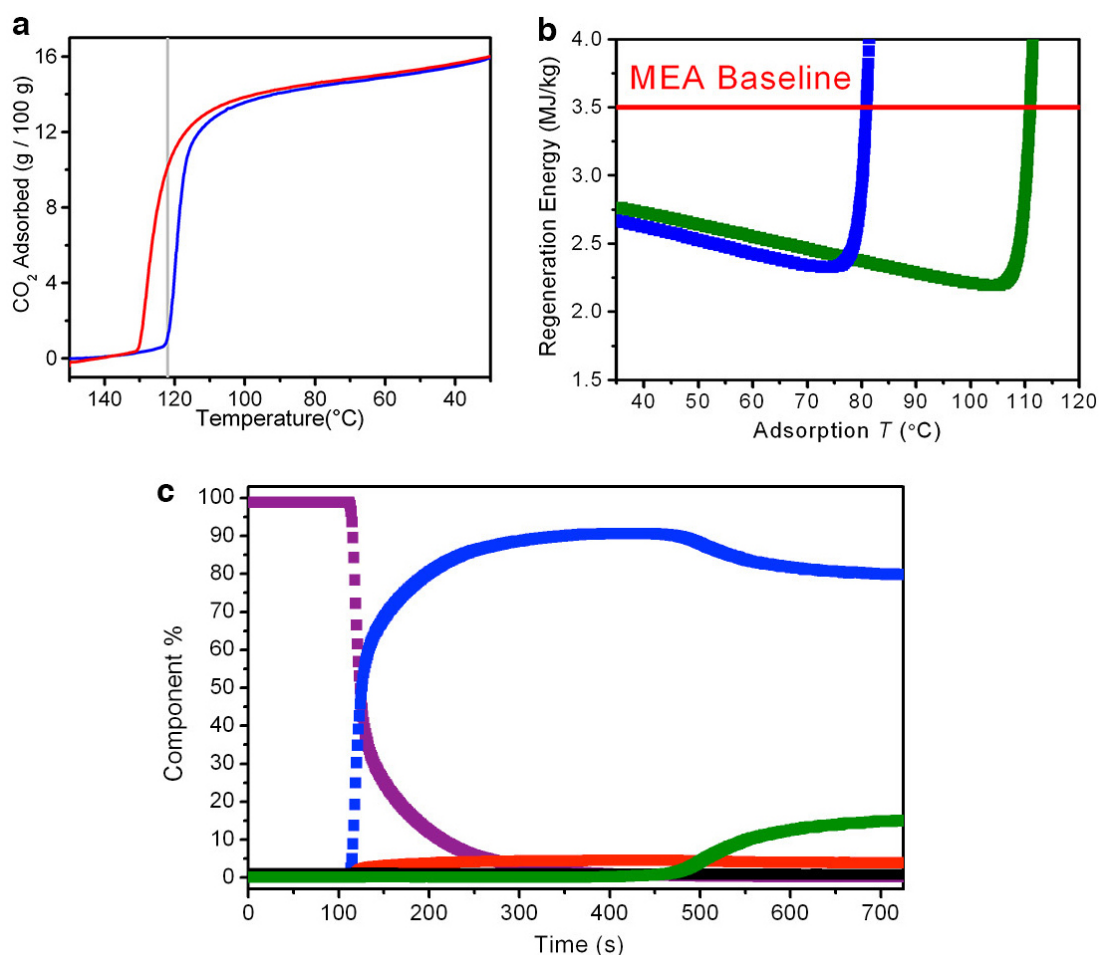
Extended Data Figure 5 | Theoretical calculations. **a**, Representation of the mapping of the hexagonal channel to a two-dimensional lattice in which each site consists of an amine that can interact with six neighbouring sites. Each amine can have one or zero CO₂ molecules adsorbed. A single site with a CO₂ adsorbed is shown in green. Pairs are allowed to form in both the crystallographic *c* direction and in the *a*-*b* plane (yellow); to model the chain mechanism, an amine in the middle of the chain is shown in red, and the amine at the end of the chain is in blue. **b**, Calculated CO₂ adsorption isotherms

indicate that only a chain model of interactions rather than pairwise adsorption interactions can give rise to the experimentally stepped isotherm shape. **c**, On the basis of calculated adsorption enthalpies, the relative positions of adsorption isotherms can be predicted from the chain model. **d**, **e**, DFT calculations reflect the experimentally observed trend that CO₂ adsorption enthalpy (**d**) is related to the strength of the nitrogen-amine bond, as reflected by the calculated metal-amine bond length (**e**).



Extended Data Figure 6 | Volumetric gas adsorption. **a**, High-pressure excess CO₂ adsorption isotherm at 25 °C for mmen-Ni₂(dobpdc) indicates that Langmuir-type adsorption behaviour is maintained even at high pressures. **b**, Isothermal adsorption measurements of water onto a sample of mmen-Mg₂(dobpdc) at 40, 75 and 100 °C. **c**, Four isothermal adsorption

measurements of CO₂ at 75 °C onto a sample of mmen-Mg₂(dobpdc) before exposure to water and after water isotherms at 40, 75 and 100 °C. No changes in the CO₂ adsorption isotherms were apparent from exposure of the sample to water.



Extended Data Figure 7 | Dynamic gas adsorption and regeneration energy.

a, Isobaric, variable temperature (ramp rate of 1 °C min⁻¹) gravimetric adsorption experiments for mmen-Mg₂(dobpdc) under 100% CO₂. Cooling from 150 to 30 °C is shown as the blue line, and heating from 30 to 150 °C as the red line. Desorption hysteresis was minimal because the phase-transition temperature and pressure are unchanged between adsorption and desorption.

b, Regeneration energy calculations for mmen-Mg₂(dobpdc) (green) and

mmen-Mn₂(dobpdc) (blue) indicate that effecting adsorption at high temperatures can be considerably more efficient than adsorption at 40 °C.

c, Transient breakthrough of 15% CO₂ (green), 4% O₂ (red), 1.5% water (black) and balance N₂ (blue) through an adsorbent bed packed with mmen-Mg₂(dobpdc) at 25 °C. The adsorbent bed was under Ar (purple) before adsorption; a breakthrough CO₂ capacity of 2.7 mmol g⁻¹ was calculated.

The fine-scale genetic structure of the British population

Stephen Leslie^{1,2,3*}, Bruce Winney^{3*}, Garrett Hellenthal^{4*}, Dan Davison⁵, Abdelhamid Boumertit³, Tammy Day³, Katarzyna Hutnik³, Ellen C. Roayrvik³, Barry Cunliffe⁶, Wellcome Trust Case Control Consortium 2†, International Multiple Sclerosis Genetics Consortium†, Daniel J. Lawson⁷, Daniel Falush⁸, Colin Freeman⁹, Matti Pirinen¹⁰, Simon Myers¹¹, Mark Robinson¹², Peter Donnelly^{9,11§} & Walter Bodmer^{3§}

Fine-scale genetic variation between human populations is interesting as a signature of historical demographic events and because of its potential for confounding disease studies. We use haplotype-based statistical methods to analyse genome-wide single nucleotide polymorphism (SNP) data from a carefully chosen geographically diverse sample of 2,039 individuals from the United Kingdom. This reveals a rich and detailed pattern of genetic differentiation with remarkable concordance between genetic clusters and geography. The regional genetic differentiation and differing patterns of shared ancestry with 6,209 individuals from across Europe carry clear signals of historical demographic events. We estimate the genetic contribution to southeastern England from Anglo-Saxon migrations to be under half, and identify the regions not carrying genetic material from these migrations. We suggest significant pre-Roman but post-Mesolithic movement into southeastern England from continental Europe, and show that in non-Saxon parts of the United Kingdom, there exist genetically differentiated subgroups rather than a general ‘Celtic’ population.

The genetic composition of human populations varies throughout the world, as a result of the interplay between population movement, admixture, natural selection and genetic drift. Characterizing such geographical population structure provides insights into demographic history and is critical to genetic studies of disease^{1–3}.

Human population structure is reasonably well understood at broad scales, for example between and within continents^{4–10}. Here we investigate structure over much finer scales, in Caucasians within the United Kingdom (UK) consisting of England, Scotland, Wales and Northern Ireland. We use ‘Britain’ (technically Great Britain) to refer to the single island consisting of modern-day England, Scotland and Wales. UK population structure has been studied before, typically on relatively small samples using various single-locus systems and recently genome-wide SNP data^{11,12}. These earlier studies show some regional variation at particular loci, with a weak, roughly north–south cline in allele frequencies genome-wide, suggesting that population structure in the UK is rather limited.

Samples and analysis

To investigate fine-scale population structure in the UK, and to provide well-characterized controls for disease studies, we assembled a sample, the People of the British Isles (PoBI) collection, as previously described¹³. Our analyses used 2,039 PoBI samples from rural areas within the UK, genotyped as part of the Wellcome Trust Case Consortium 2 (WTCCC2), who had all four grandparents born within 80 km of each other. We thus effectively sample DNA from the grandparents. The grandparents’ average year of birth was 1885 (s.d. 18 years). As the DNA from each PoBI participant is a random sample of their grandparents’ DNA, our approach allows investigation of fine-scale

population structure in rural areas of the UK before the major population movements of the twentieth century.

To provide context for the UK samples, we analysed 6,209 samples from 10 countries in continental Europe genotyped in the WTCCC2 study of multiple sclerosis¹⁴. To ensure compatibility between the PoBI and continental European samples we restricted attention to autosomal SNPs genotyped in both samples (approximately 500,000 SNPs, see Methods).

Fine-scale UK population differentiation

Consistent with earlier studies of the UK, population structure within the PoBI collection is very limited. The average of the pairwise F_{ST} estimates between each of the 30 sample collection districts is 0.0007, with a maximum of 0.003 (Supplementary Table 1).

Against this background of very limited structure within the UK, we applied a recently developed method for detecting fine-scale population structure, fineSTRUCTURE¹⁵, to the PoBI samples, to look for more subtle effects. See Methods (also Extended Data Figs 1 and 2) for an informal description, details, interpretation under both discrete and isolation-by-distance models, assessment of convergence, and enhancements to the algorithm as applied in this study. In contrast to commonly used approaches such as principal components or ADMIXTURE¹⁶, fineSTRUCTURE explicitly models the correlation between nearby SNPs and uses extended multi-marker haplotypes throughout the genome. This substantially increases its power to detect subtle levels of genetic differentiation.

The fineSTRUCTURE algorithm can divide samples into genetic clusters hierarchically, from coarser to finer levels of structuring. We

¹Murdoch Childrens Research Institute, Royal Children’s Hospital, Flemington Road, Parkville, Victoria 3052, Australia. ²University of Melbourne, Department of Mathematics and Statistics, Parkville, Victoria 3010, Australia. ³University of Oxford, Department of Oncology, Old Road Campus Research Building, Roosevelt Drive, Oxford OX3 7DQ, UK. ⁴University College London Genetics Institute, Darwin Building, Gower Street, London WC1E 6BT, UK. ⁵Counsyl, 180 Kimball Way, South San Francisco, California 94080, USA. ⁶University of Oxford, Institute of Archaeology, 36 Beaumont Street, Oxford OX1 2PG, UK. ⁷University of Bristol, Department of Mathematics, University Walk, Bristol BS8 1TW, UK. ⁸College of Medicine, Swansea University, Singleton Park, Swansea SA2 8PP, UK. ⁹The Wellcome Trust Centre for Human Genetics, Roosevelt Drive, Oxford OX3 7BN, UK. ¹⁰University of Helsinki, P.O. Box 20, Helsinki, FI-00014, Finland. ¹¹University of Oxford, Department of Statistics, 1 South Parks Road, Oxford OX1 3TG, UK. ¹²University of Oxford, University Museum of Natural History, Parks Road, Oxford OX1 3PW, UK.

*These authors contributed equally to this work.

§These authors jointly supervised this work.

†Information about participants appears in the Supplementary Information.

applied fineSTRUCTURE to the PoBI samples' genetic data without reference to the known geographical locations. The genetic clustering can be assessed with respect to geography by plotting individuals on a map of the UK (at the centroid of their grandparents' places of birth) and examining the inferred genetic clusters, for different levels of the hierarchical clustering.

Figure 1 shows this map for 17 clusters, together with the tree showing how these clusters are related at coarser levels of the hierarchy. (There is nothing special about this level of clustering, but it is convenient for describing some of the main features of our analysis; Supplementary Fig. 1 depicts maps showing other levels of the hierarchical clustering.) The correspondence between the genetic clusters and geography is striking: most of the genetic clusters are highly localized, with many occupying non-overlapping regions. Because the genetic clustering made no reference to the geographical location of the samples, the resulting correspondence between genetic clusters and geography reassures us that our approach is detecting real population differentiation at fine scales. Our approach can separate groups in close proximity, such as in Cornwall and Devon in southwest England, where the genetic clusters closely match the modern county boundaries, or in Orkney, off the north coast of Scotland.

It is instructive to consider the tree that describes the hierarchical splitting of the 2,039 genotyped individuals into successively finer clusters (Fig. 1). The coarsest level of genetic differentiation (that is, the assignment into two clusters) separates the samples in Orkney from all others. Next the Welsh samples separate from the other non-Orkney samples. Subsequent splits reveal more subtle differentiation (reflected in the shorter distances between branches), including separation of north and south Wales, then separation of the north of England, Scotland and Northern Ireland from the rest of England, and separation of samples in Cornwall from the large English cluster. There is a single large cluster (red squares) that covers most of central and southern England and extends up the east coast. Notably, even at the finest level of differentiation returned by fineSTRUCTURE (53 clusters), this cluster remains largely intact and contains almost half the individuals (1,006) in our study.

Although larger than between the sampling locations, estimated F_{ST} values between the clusters represented in Fig. 1 are small (average 0.002, maximum 0.007, Supplementary Table 2), confirming that differentiation is subtle. On the other hand, all comparisons between pairs of clusters of their patterns of ancestry as estimated by fineSTRUCTURE show highly significant differences (Supplementary Table 3).

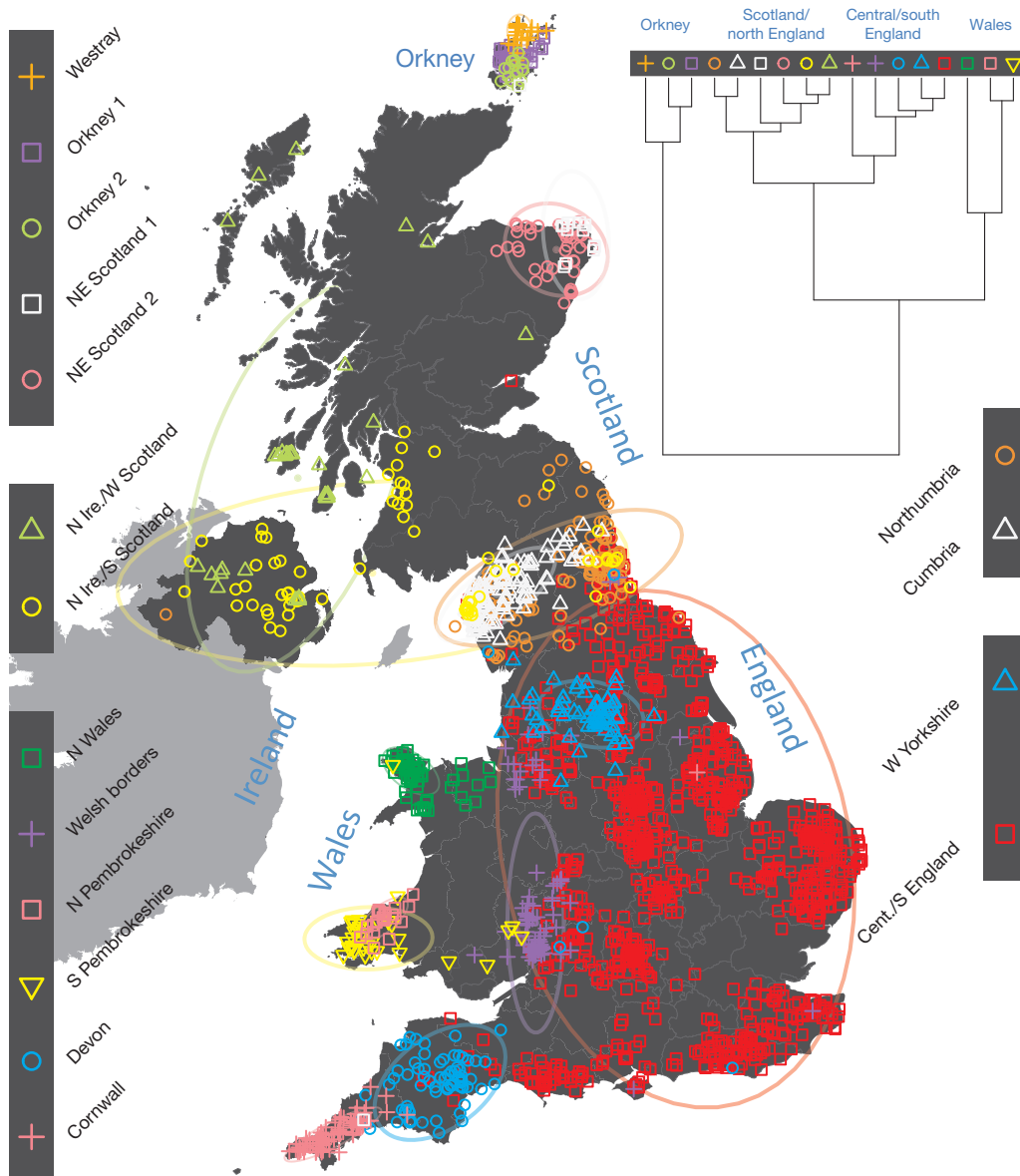


Figure 1 | Clustering of the 2,039 UK individuals into 17 clusters based only on genetic data. For each individual, the coloured symbol representing the genetic cluster to which the individual is assigned is plotted at the centroid of their grandparents' birthplaces. Cluster names are in side-bars and ellipses give an informal sense of the range of each cluster (see Methods). No relationship between clusters is implied by the colours/symbols. The tree (top right) depicts the order of the hierarchical merging of clusters (see Methods for the interpretation of branch lengths). Contains OS data © Crown copyright and database right 2012. © EuroGeographics for some administrative boundaries.

We compared our approach to two widely used analysis tools, namely principal components^{4,12,17} and ADMIXTURE¹⁶ (Extended Data Fig. 3). Both approaches broadly separate samples from Wales and from Orkney, but are not able to distinguish many of the other clusters found by fineSTRUCTURE. We also performed analyses to show that the clustering is not an artefact of our sampling scheme preferentially selecting related individuals (see Methods, Extended Data Fig. 4 and Supplementary Note).

UK clusters in relation to Europe

Genetic differences between UK clusters might in part reflect their relative isolation from each other, and in part differing patterns of migration and admixture from populations outside the UK. To gain further insight into this second aspect, we first applied similar fineSTRUCTURE analyses to 6,209 samples from continental Europe (henceforth referred to as ‘Europe’, see Extended Data Fig. 5a for the distribution of the samples by region), and then characterized the genetic composition of the UK clusters with respect to the genetic groups we found in Europe. A fuller analysis of the clustering within Europe and its interpretation will be described elsewhere.

To avoid confusion below, we will refer to each of the 17 sets of individuals defined by our fineSTRUCTURE analyses in the UK as a ‘cluster’, and to each of the sets of individuals defined in our analyses of Europe as a ‘group’. We focus in these analyses on the division of the European samples into 51 such groups (Extended Data Fig. 5b). We italicise names of UK clusters, to distinguish them from the geographic region (for example, the pink cross cluster *Cornwall*, and the county Cornwall). European groups are each given a unique identifying number (these are consecutive at the finest level of clustering, but not at the level we consider). In the text, groups are identified by this number and, for clarity, a three-letter label identifying the country (or countries) where the group is mainly represented (for example, GER6 for the group labelled ‘6’, which is mostly found in Germany).

For each UK cluster we estimated an ‘ancestry profile’ which characterizes the ancestry of the cluster as a mixture of the ancestry of the 51 European groups. (see Methods for details, also Supplementary Table 4). As for the fineSTRUCTURE clustering, these analyses use no geographical information. The estimated ancestry profiles are illustrated in Fig. 2 which also depicts the sampling locations in Europe of the groups contributing to the ancestry profiles (see also Extended Data Fig. 6a). Note that it is possible for distinct clusters within the UK to have very similar ancestry profiles: for example, two UK regions could receive similar contributions from a set of European groups (thus similar ancestry profiles) but then evolve separately (leading to different patterns of shared ancestry within and between the regions, and thus to distinct clusters in fineSTRUCTURE).

The bar charts in Fig. 2 show that some European groups feature substantially in the ancestry profiles of all UK clusters. These are: GER6 (yellow green) found predominantly in western Germany; BEL11 (green), in the northern, Flemish, part of Belgium; FRA14 (light blue), in north-west France; DEN18 (dark blue), in Denmark; SFS31 (blue/purple) in southern France and Spain. In contrast, some European groups feature substantially in the ancestry profiles of some UK clusters but are absent from those of other UK clusters: GER3 (orange), in northern Germany; FRA12 (dark green), in France; and FRA17 (blue), also in France. Two Swedish groups (SWE117 and SWE121) feature in the ancestry profiles of the UK clusters, with Norwegian groups (shades of purple) featuring substantially in the ancestry profiles of the Orkney clusters, and to a lesser extent the clusters involving Scotland and Northern Ireland.

Discussion

The application of powerful haplotype based analysis methods to genome-wide SNP data from a large, carefully-collected, UK sample reveals a rich pattern of subtle fine-scale genetic differentiation within the UK, which shows a marked concordance with geography. Few of these details have been captured previously.

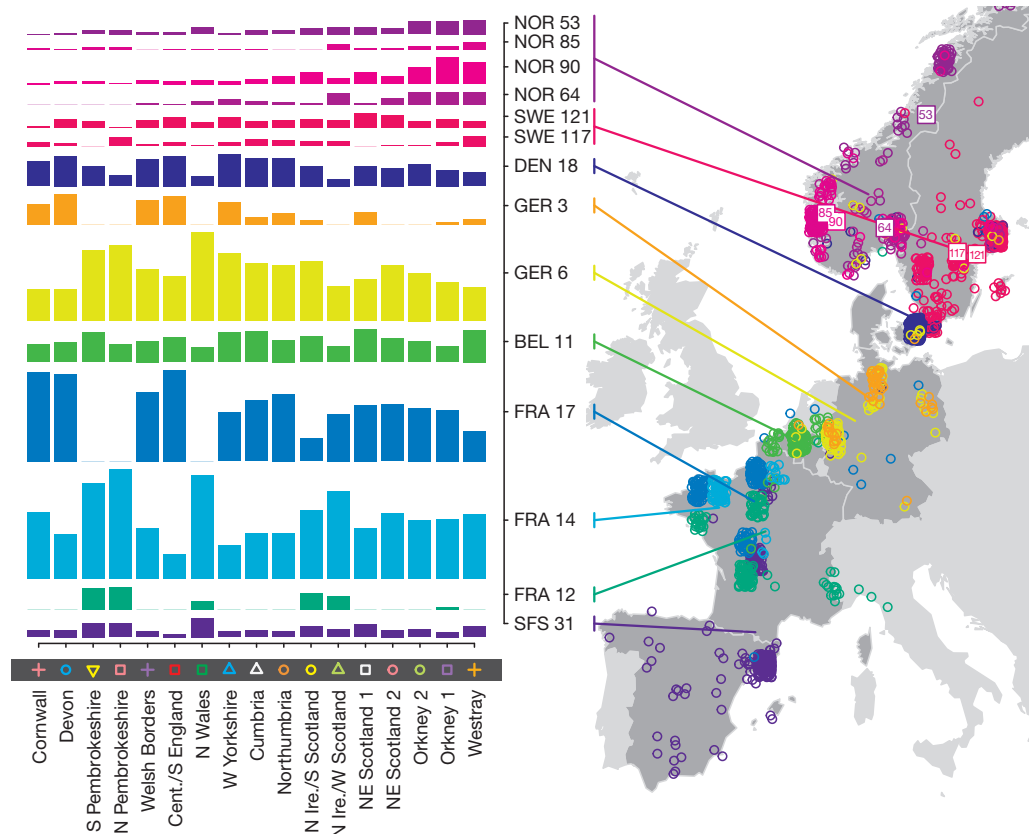


Figure 2 | European ancestry profiles for the 17 UK clusters. Each row represents one of the 51 European groups (labels at right) that were inferred by clustering the 6,029 European samples using fineSTRUCTURE. Only European groups that make at least 2.5% contribution to the ancestry profile of at least one UK cluster are shown. Each column represents a UK cluster. Coloured bars have heights representing the proportion of the UK cluster's ancestry best represented by that of the European group labelled with that colour. The map shows the location (when known at regional level) of the samples assigned to each European group (some sample locations are jittered and/or moved for clarity, see Methods). Lines join group labels to the centroid of the group, or collection of groups (Norway, Sweden, with individual group centroids marked by group number). © EuroGeographics for the administrative boundaries.

The clustering (Fig. 1 and Supplementary Fig. 1) is notable both for its exquisite differentiation over small distances and the stability of some clusters over very large distances. Genetic differentiation within the UK is not related in a simple way to geographical distance. Examples of fine-scale differentiation include the separation of: islands within Orkney; Devon from Cornwall; and the Welsh/English borders from surrounding areas. The edges between clusters follow natural geographical boundaries in some instances, for example, between *Devon* and *Cornwall* (boundaries the Tamar Estuary and Bodmin Moor), and Orkney is separated by sea from Scotland. However, in many instances clusters span geographic boundaries; for example, the clusters in Northern Ireland span the sea to Scotland.

Although the branch lengths of the hierarchical clustering tree in Fig. 1 are not easy to interpret directly, they are indicative of the relative differentiation between UK clusters, so that for example, the differences between Orkney, Wales and the remainder of the UK are substantial compared to some of the finer differences (splits closer to the tips of the tree). North and south Wales are about as distinct genetically from

each other as are central and southern England from northern England and Scotland, and the genetic differences between Cornwall and Devon are comparable to or greater than those between northern English and Scottish samples, and to those between islands in Orkney.

To facilitate further discussion, Fig. 3 and Extended Data Fig. 7 give an overview of the major population groups and movements of people within and into the UK at different times, based on archaeological, historical and linguistic evidence. For more detail see the Supplementary Note.

Our observation that samples in Orkney differ genetically from those in the rest of the UK has been noted before^{18–21} and is consistent with the historical settlement, and long-term control of Orkney by Norse Vikings (Orkney was a part of Norway from 875 to 1472). Further, the estimated ancestry profiles of the Orkney clusters show substantial contributions from groups in Norway (Fig. 2). This consistency with history and archaeology provides external validation of our approach.

Our approach is clearly powered to detect quite subtle levels of population structure. Not finding such structure in central and southern England is thus informative. Although some structure may exist within

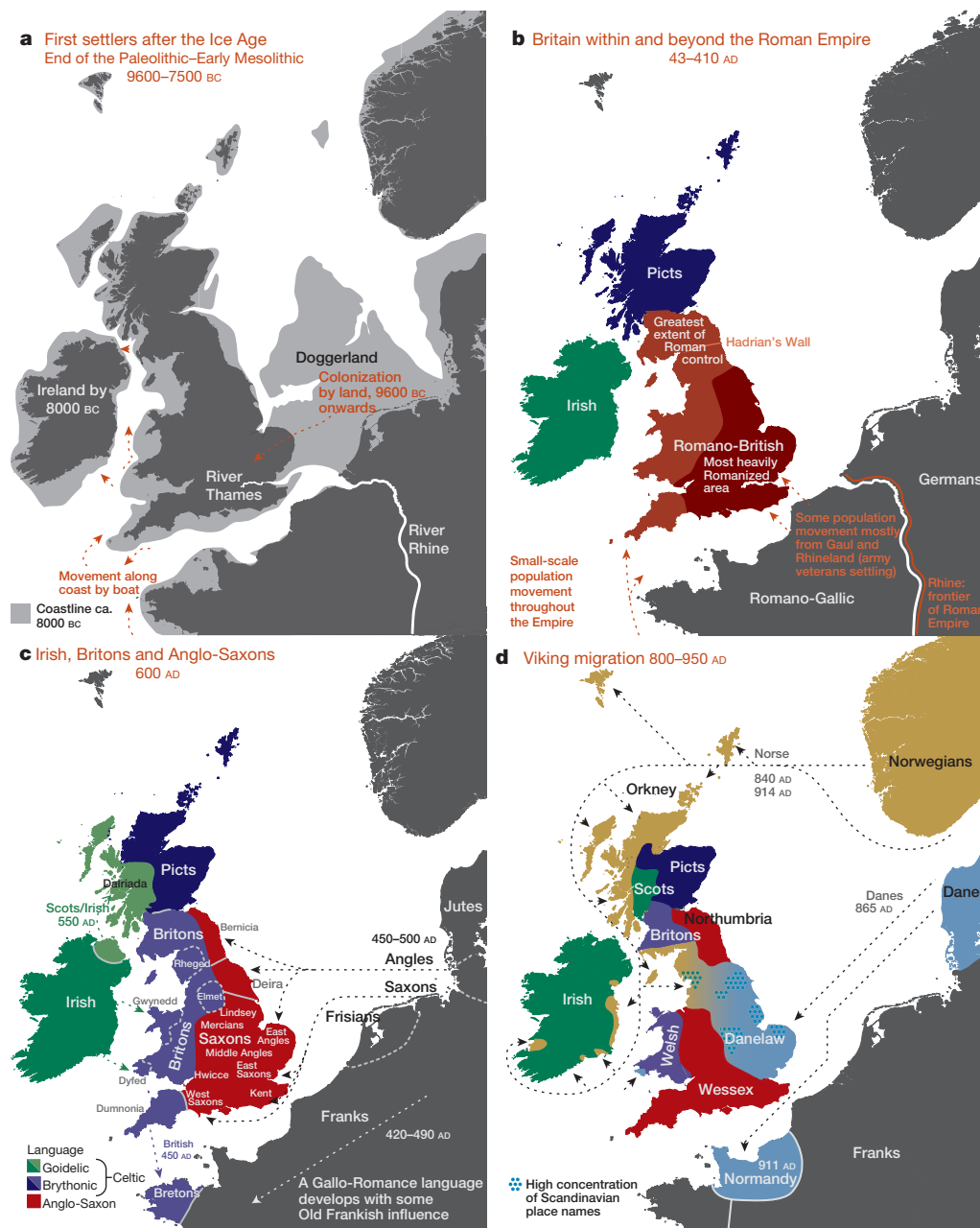


Figure 3 | Major events in the peopling of the British Isles. See Supplementary Note for further details. **a**, The routes taken by the first settlers after the last ice age. **b**, Britain during the period of Roman rule. **c**, The regions of ancient British, Irish and Saxon control. **d**, The migrations of Norse and Danish Vikings. The main regions of Norse Viking (light brown) and Danish Viking (light blue) settlement are shown. © EuroGeographics for the administrative boundaries (coastlines).

this region, there must have been sufficient movement of people, and hence of their DNA, since the last major invasions of the UK to make it relatively homogeneous genetically. This does not require large-scale population movements; it could be achieved by relatively local migration over many generations. This region of Britain lacks major geographical and (for the most part since the Roman occupation) geo-political barriers to human movement.

Other UK clusters may well reflect historical events. For example, several genetic clusters in Fig. 1 match the geo-political boundaries in Fig. 3c, and may represent remnants of communities/kingdoms present after the Saxon migrations, while the cluster spanning Northern Ireland and southern Scotland may reflect the 'Ulster plantations'. The Supplementary Note contains further observations relating to the genetic clustering.

Relative isolation has clearly been a major determinant of fine-scale population structure within the UK. To assess the role of a different possible cause, namely differential migration into different parts of the UK, we estimated European ancestry profiles for each of the UK genetic clusters (Fig. 2). Here we must use modern-day groupings, in Europe and the UK, as surrogates for the sources and results of major migration events. Population movements between these events and the present, involving either the source populations or recipient groups, will attenuate signals of the original migration. For this and other reasons, it is hard to provide definitive explanations for our observations. Nonetheless, genetic differences persist through many generations and where we can check our conclusions against historical evidence, there is good concordance. In what follows we focus on the most likely explanations for various observations. See Supplementary Note for a fuller discussion. For definiteness, we focus on the clustering in Fig. 1 and Extended Data Fig. 5b, although other levels are informative. Analysis of additional UK and European samples, particularly in regions where our data are sparse (for example, central Wales and Scotland, Spain, the Netherlands) would improve our ability to infer historical events.

The observation (Fig. 2 and Supplementary Table 4) that particular European groups (for example, GER3, FRA12, FRA17) contribute substantially to the ancestry profiles of some, but not all, UK clusters strongly suggests that at least some of the structure we observe in the UK results from differential input of DNA to different parts of the UK: the absence in particular UK clusters of ancestry from specific European groups is best explained by the DNA from those European groups never reaching those UK clusters. A critical observation which follows is that groups which contribute significantly to the ancestry profiles of all UK clusters most probably represent, at least in part, migration events into the UK that are relatively old, since their DNA had time to spread throughout the UK. Conversely, groups that contribute to the ancestry profiles of only some UK clusters most probably represent more recent migration events, with the resulting DNA not yet spread throughout the UK by internal migration. 'Old' and 'recent' here are relative terms—we can infer the order of some events in this way but not their absolute times. Although we refer to migration events, we cannot distinguish between movements of reasonable numbers of people over a short time or ongoing movements of smaller numbers over longer periods.

Applying this approach suggests a relative ordering of the peopling of the British Isles. For a full discussion, and caveats, see Supplementary Note. Briefly, the earliest migrations whose descendants survive to make a substantial contribution to the present population are best captured by three groups in our European analyses, GER6 (western Germany), BEL11 (Belgium), and FRA14 (north-western France). These groups still contribute to current patterns of population differentiation (Fig. 2, see also Extended Data Fig. 6). Other European groups may reflect early migrations into the UK, but with smaller contribution, including SFS31 (southern France/Spain), at least part of DEN18 (Denmark), and possibly parts of Norway and Sweden. A subsequent migration, best captured by FRA17 (France), contributed a substantial amount of ancestry to the UK outside Wales. Although we cannot formally exclude this being part of the Saxon migration, this seems unlikely (see Methods)

and instead it might represent movement of people taking place between the early migrations and those known from historical records. Migrations represented by FRA12 essentially only affect Wales and Northern Ireland and/or Scotland. We also see clear signals of some of the known historical migrations and settlements, including the Saxons (GER3, northern Germany, and probably much of DEN18, Denmark) and the Norse Vikings (NOR53–NOR90).

To further shed light on two major migration events, in Orkney and in central and southern England respectively, we applied a distinct analytical tool, GLOBETROTTER²² (Extended Data Figs 8 and 9). Informally, GLOBETROTTER exploits information in the rate of decay of shared haplotype segments to test for the presence of recent admixture, to identify groups contributing, and then date the admixture.

GLOBETROTTER detected strong evidence ($P < 0.01$) that the largest Orkney cluster (*Orkney 1*) was influenced by a recent admixture event with an overall contribution of ~25% of the DNA from groups in Norway, confirming that the Norwegian contribution in the ancestry profile for this cluster reflects recent admixture (Extended Data Fig. 9). The approach assumed the simplest model (a single pulse of admixture), and estimated this to have occurred 29 generations ago (95% confidence interval (CI): 18–39 generations), corresponding to year 1100 (95% CI: 830–1418), assuming a 28 year generation time²²; no clear evidence was found of multiple admixture dates. We expect less precise estimates for the other two Orkney clusters (due to their smaller sample size), but these were consistent with those for *Orkney 1*. For *Cent./S England* the method also detected an admixture event, with a contribution of ~35% of DNA from GER3, the group in north-western Germany, and an estimated date of 38 generations (95% CI: 36–40 generations), corresponding to year 858 (95% CI: 802–914) (Extended Data Fig. 9). The GLOBETROTTER analyses detect likely source populations for the known historical migrations (Norse Vikings and Saxons, respectively) with the estimated proportion contributed by these sources close to that estimated in the ancestry profiles. Note that a migration event is likely to precede any subsequent population admixture, possibly substantially so, if the migrants mate largely within the migrant group for some time after their migration. Further, admixture is likely to be a gradual process, so that using a model of a single pulse of admixture in GLOBETROTTER is likely to estimate a time after the commencement of admixture. For these reasons, the admixture dates estimated by GLOBETROTTER should provide upper bounds on the dates of the migrations²², as for both examples here, where the estimated dates are 200 or more years after the known dates of the migrations, suggesting that the mixing was indeed a gradual process.

After the Saxon migrations, the language, place names, cereal crops and pottery styles all changed from that of the existing (Romano-British) population to those of the Saxon migrants. There has been ongoing historical and archaeological controversy about the extent to which the Saxons replaced the existing Romano-British populations. Earlier genetic analyses, based on limited samples and specific loci, gave conflicting results. With genome-wide data we can resolve this debate. Two separate analyses (ancestry profiles and GLOBETROTTER) show clear evidence in modern England of the Saxon migration, but each limits the proportion of Saxon ancestry, clearly excluding the possibility of long-term Saxon replacement. We estimate the proportion of Saxon ancestry in *Cent./S England* as very likely to be under 50%, and most likely in the range of 10–40%.

A more general conclusion of our analyses is that while many of the historical migration events leave signals in our data, they have had a smaller effect on the genetic composition of UK populations than has sometimes been argued. In particular, we see no clear genetic evidence of the Danish Viking occupation and control of a large part of England, either in separate UK clusters in that region, or in estimated ancestry profiles, suggesting a relatively limited input of DNA from the Danish Vikings and subsequent mixing with nearby regions, and clear evidence for only a minority Norse contribution (about 25%) to the current Orkney population.

We saw no evidence of a general 'Celtic' population in non-Saxon parts of the UK. Instead there were many distinct genetic clusters in these regions, some amongst the most different in our study, in the sense of being most separated in the hierarchical clustering tree in Fig. 1. Further, the ancestry profile of *Cornwall* (perhaps expected to resemble other Celtic clusters) is quite different from that of the Welsh clusters, and much closer to that of *Devon*, and *Cent./S England*. However, the data do suggest that the Welsh clusters represent populations that are more similar to the early post-Ice-Age settlers of Britain than those from elsewhere in the UK.

In summary, we have presented the first (to our knowledge) fine-scale dissection of subtle levels of genetic differentiation within a country, by using careful sampling, genomic data and powerful statistical methods. The resulting genetic clusters, and the characterization of their ancestry in terms of European groups, provide important and novel insights into the peopling of the British Isles.

Genetic information can augment archaeological, linguistic and historical approaches to understanding population history. It also complements them, in providing evidence relating to the bulk of ordinary people rather than the successful elite. We hope that our study will act as a proof-of-principle for the power of such detailed genetic analyses.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 November 2013; accepted 13 January 2015.

- Cardon, L. R. & Bell, J. I. Association study designs for complex diseases. *Nature Rev. Genet.* **2**, 91–99 (2001).
- Marchini, J., Cardon, L. R., Phillips, M. S. & Donnelly, P. The effects of human population structure on large genetic association studies. *Nature Genet.* **36**, 512–517 (2004).
- Bodmer, W. & Bonilla, C. Common and rare variants in multifactorial susceptibility to common diseases. *Nature Genet.* **40**, 695–701 (2008).
- Cavalli-Sforza, L. L., Menozzi, P. & Piazza, A. *The History and Geography of Human Genes* (Princeton Univ. Press, 1994).
- Quintana-Murci, L. *et al.* Genetic evidence of an early exit of *Homo sapiens sapiens* from Africa through eastern Africa. *Nature Genet.* **23**, 437–441 (1999).
- Conrad, D. F. *et al.* A worldwide survey of haplotype variation and linkage disequilibrium in the human genome. *Nature Genet.* **38**, 1251–1260 (2006).
- The 1000 Genomes Project Consortium. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
- Botigué, L. R. *et al.* Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc. Natl Acad. Sci. USA* **110**, 11791–11796 (2013).
- Ralph, P. & Coop, G. The geography of recent genetic ancestry across Europe. *PLoS Biol.* **11**, e1001555 (2013).
- Hellenthal, G., Auton, A. & Falush, D. Inferring human colonization history using a copying model. *PLoS Genet.* **4**, e1000078 (2008).
- The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
- O'Dushlaine, C. T. *et al.* Population structure and genome-wide patterns of variation in Ireland and Britain. *Eur. J. Hum. Genet.* **18**, 1248–1254 (2010).
- Winney, B. *et al.* People of the British Isles: preliminary analysis of genotypes and surnames in a UK-control population. *Eur. J. Hum. Genet.* **20**, 203–210 (2012).
- The International Multiple Sclerosis Genetics Consortium & The Wellcome Trust Case Control Consortium 2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214–219 (2011).
- Lawson, D. J., Hellenthal, G., Myers, S. & Falush, D. Inference of population structure using dense haplotype data. *PLoS Genet.* **8**, e1002453 (2012).
- Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).
- Pirinen, M., Donnelly, P. & Spencer, C. C. A. Efficient computation with a linear mixed model on large-scale data sets with applications to genetic studies. *Ann. Appl. Stat.* **7**, 369–390 (2013).
- Wilson, J. F. *et al.* Genetic evidence for different male and female roles during cultural transitions in the British Isles. *Proc. Natl Acad. Sci. USA* **98**, 5078–5083 (2001).
- Capelli, C. *et al.* A Y chromosome census of the British Isles. *Curr. Biol.* **13**, 979–984 (2003).
- Goodacre, S. *et al.* Genetic evidence for a family-based Scandinavian settlement of Shetland and Orkney during the Viking periods. *Heredity* **95**, 129–135 (2005).
- Wells, R. S. *et al.* The Eurasian heartland: a continental perspective on Y-chromosome diversity. *Proc. Natl Acad. Sci. USA* **98**, 10244–10249 (2001).
- Hellenthal, G. *et al.* A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Cheshire for his advice. We thank the UK Office for National Statistics, the National Records of Scotland, and the Northern Ireland Statistics and Research Agency for providing the boundaries used for the UK maps. We note that census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland. We further acknowledge the provision of maps from Eurostat, which are copyright EuroGeographics for the administrative boundaries. We acknowledge support from the Wellcome Trust (072974/Z/03/Z, 088262/Z/09/Z, 075491/Z/04/Z, 075491/Z/04/A, 075491/Z/04/B, 090532/Z/09/Z, 084818/Z/08/Z, 095552/Z/11/Z, 085475/Z/08/Z, 098387/Z/12/Z, 098386/Z/12/Z), the Academy of Finland (257654) and the Australian National Health and Medical Research Council (APP1053756). P.D. was supported in part by a Wolfson-Royal Society Merit Award.

Author Contributions W.B. conceived and directed the PoBI project. P.D. directed the analysis and sample genotyping. B.W., A.B., T.D., K.H., E.C.R. and W.B. collected the UK (PoBI) samples and extracted DNA. IMSCG provided the European samples' genotypes and geographical information. Sample genotyping and quality control was performed by WTCCC2 for both the UK and European genotype data. S.L., G.H., S.M. and P.D. performed the major analyses with contributions from B.W., D.D., D.J.L., D.F., C.F., M.R., M.P. and W.B. M.R. and B.C. provided historical and archaeological information and context. G.H. made Extended Data Fig. 2. S.L. produced all the other figures. P.D., S.L., B.W., G.H., S.M., M.R. and W.B. wrote the manuscript. All authors reviewed the manuscript.

Author Information Genotype data, as well as location information at county level (or aggregated across counties where there are small numbers of samples associated with a particular county), will be made available by the WTCCC access process, via the European Genotype Archive (<https://www.ebi.ac.uk/ega/>) under accession numbers EGAS00001000672 and EGAD00010000632. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.D. (donnelly@well.ox.ac.uk).

METHODS

Samples, genotyping and QC. The sampling scheme and general information about the UK sample is described elsewhere¹³. Briefly, the aim was to collect samples from rural regions of the UK, for whom all four grandparents were born close to each other. In total 4,371 samples were collected as part of the PoBI project. Of these 2,886 were genotyped on the Illumina Human 1.2M-Duo genotyping chip as part of the Wellcome Trust Case Control Consortium 2 (WTCCC2) studies, with 2,510 passing the WTCCC2 genotype quality control (QC) procedures²³. We then applied a geographic filter, which imposed a maximum pairwise distance between each sample's grandparents' places of birth of 80 km, leaving 2,039 samples available for analysis. In what follows we refer to these samples as the 'UK sample(s)'. We give a detailed description of the choice of SNPs used for our analyses below.

For the European ancestry profile analysis we used 6,209 samples from the WTCCC2 multiple sclerosis study¹⁴, of which 5,682 were cases and 527 were controls. We excluded all samples from the UK and Ireland (see 'treatment of Eire' below). Extended Data Fig. 5a shows a breakdown of sample numbers by region. In the following text we refer to these continental European samples as the 'European sample(s)'. The European samples were genotyped on the Illumina Human 660-Quad chip as previously described¹⁴. These samples had already passed through the WTCCC2 SNP and sample quality control procedures¹⁴.

For all analyses we used intersections of the autosomal SNPs available for the UK and European data sets, constructed in the following manner: we excluded SNPs in the HLA region, and, for analyses involving the European samples, SNPs in major multiple sclerosis associated regions (although any effect of the use of disease samples should be small in analyses of genome-wide data). More specifically, we first took the full intersection of the UK and European data SNP sets. We removed a 15 Mb region surrounding the HLA region on chromosome 6 because the European samples were comprised of multiple sclerosis case samples, a disease with strong HLA associations. This left 575,236 SNPs that were transferred to the haplotype inference (phasing) step (see next section). Within the phasing software (IMPUTE2) further SNPs were excluded based on WTCCC2 quality control procedures, which—in addition to IMPUTE2's internal removal of SNPs due to strand issues or lack of overlap between the SNP array and the reference panel haplotypes—removed 15,211 of these SNPs before phasing. After phasing, SNPs with IMPUTE2 info-threshold ≤ 0.975 and SNPs that were singletons among all phased data were removed (these data include all POBI and European samples). This left 524,699 SNPs. For the analyses using only the UK data (the clustering analysis labelled 'analysis A' in the next section) 522,862 SNPs were used in the CHROMOPAINTER/fineSTRUCTURE analyses (see next section), as the rest were monomorphic in the UK set of 2,039 individuals. For further analyses, using the European data (labelled 'analysis B' and 'analysis C' in the next section), multiple sclerosis associated SNPs (regions defined by linkage disequilibrium around major loci of suggestive association with multiple sclerosis) were removed, as well as some other SNPs for technical reasons. In total this removed SNPs from 56.8 Mb of the genome. This resulted in 515,981 SNPs remaining for the analyses involving European samples. In summary, there were 522,862 SNPs available for the UK clustering analyses, and 515,981 SNPs available for the analyses involving European samples. A complete list of rsIDs is available at (<http://www.well.ox.ac.uk/POBI>).

Inference of population structure. To aid in understanding we give an informal description of the approach we applied for inferring fine-scale population structure. This is followed by a more detailed elaboration of our analysis. A critical feature of the algorithm, unlike other common approaches to detecting population structure such as principal components, ADMIXTURE¹⁶ or STRUCTURE²⁴, is that it explicitly models the correlation structure amongst nearby SNPs due to linkage disequilibrium, making use of the information in extended multi-marker haplotypes throughout the genome. This adds substantially to fineSTRUCTURE's power to detect subtle levels of genetic differentiation. It has been known since the early HLA studies that methods that account for linkage disequilibrium are more informative for studies of human population structure than approaches which treat each locus marginally²⁵.

Very informally, in the fineSTRUCTURE approach, haplotype phase was first inferred in each sample, after which each resulting haploid genome is broken into pieces, in such a way that for each piece the method identifies the homologous piece in another individual to which it is most similar. This can be thought of as identifying the other individual in the collection with the most similar ancestry for that part of the genome (the average size of these pieces varies across individuals, but has median 0.51 cM with IQR 0.44–0.63 cM). For each individual, one can tally up the number of pieces over which its genome is closest to each other sampled individual. These individual vectors of similarity counts are then used to cluster together individuals with similar ancestries, using a model-based statistical algorithm (fineSTRUCTURE) fitted by Markov chain Monte Carlo. The choice as to the number of clusters, and the assignment of individuals to clusters, is made so as to maximise the posterior probability under the probability model used for

clustering in fineSTRUCTURE. In the PoBI analysis, this yields 53 clusters of individuals. Similar clusters are then merged hierarchically to give a tree which can be used to describe population structure at different levels of granularity, as we describe below.

More formally, haplotypes were inferred (phased) jointly for all individuals used in the study (that is, the UK and European samples) with IMPUTE2²⁶, using the default values (see http://mathgen.stats.ox.ac.uk/impute/impute_v2.html#mcmc_options). The reference data used are available from the IMPUTE2 website (http://mathgen.stats.ox.ac.uk/impute/impute_v2.html#reference).

Next, we used the algorithm implemented in the CHROMOPAINTER program¹⁵ to represent the DNA of individuals as mosaics of the DNA from other individuals.

We performed three separate CHROMOPAINTER analyses:

- Form each haplotype of a UK individual as a mosaic of all UK haplotypes excluding those of that individual.
- Form each UK haplotype as a mosaic of all European haplotypes.
- Form each haplotype of a European individual as a mosaic of all European haplotypes excluding those of that individual.

For each analysis, A–C, we ran the algorithm implemented in CHROMOPAINTER as recommended by the authors, except for a minor change to the value of a single parameter for analysis A, implemented for technical reasons. Specifically, we initially applied CHROMOPAINTER to a subset of individuals and chromosomes (chosen as described below) using 10 iterations of its expectation-maximization (EM) algorithm to infer the genome-wide average switch and global emission rates in CHROMOPAINTER's Hidden Markov model. We averaged the inferred values of each across the chromosomes and individuals used, weighting chromosomes by their relative size, and fixed these final switch and global emission rates in a final run of CHROMOPAINTER on all individuals and chromosomes. This final CHROMOPAINTER run gave the final 'counts' and 'lengths' values used in all subsequent analyses. For analysis A, we inferred switch and global emission rates averaging across chromosomes 4, 10, 15, 22 (using weights of 187, 131, 81 and 34, respectively) and 20 individuals from each of 30 UK sample regions (counties or districts from which the PoBI samples were collected, from across the whole UK), starting with an initial switch rate of $400,000/(2N_{UK})$, where N_{UK} is the number of samples used for the UK analyses, and a default emission rate. For analyses B and C, we inferred switch and global emission rates averaging across chromosomes 1, 8, 15, 22 (using weights of 219, 142, 81 and 34, respectively) for 20 individuals from each of 30 United Kingdom regions, and 20 individuals out of every 200 in a combined file of all European subjects, starting with an initial switch rate of $400,000/(2N_E)$, where N_E is the number of samples used for the European analyses, and a default emission rate. Previous work with CHROMOPAINTER has shown that deviations of the switch rate (even up to a factor of 10) have little effect on CHROMOPAINTER's inference (data not shown). Finally, for analysis C, we set the expected number of haplotypic segments to define a region (that is, the 'k' switch) to CHROMOPAINTER's default value of 100 in order to estimate a normalization parameter (denoted by 'c') subsequently used by the clustering program fineSTRUCTURE¹⁵. In contrast, we set this value to 50 (that is, using 'k 50') for analysis A. This slight deviation from CHROMOPAINTER's default value was implemented for analysis A because some UK individuals shared relatively long haplotype segments with other UK haplotypes, such that they did not always have 100 total such segments across the entirety of some of the smaller chromosomes. We used the June 2008 build 36 genetic map from the HapMap webpage (http://hapmap.ncbi.nlm.nih.gov/downloads/recombination/2008-03_rel22_B36/rates/).

CHROMOPAINTER provides estimates of the counts of haplotype segments and total length of DNA (in cM) for which an individual shares most recent common ancestry with a set of other individuals. When summed across all 22 autosomes we refer to the vector of these counts as the 'copying profile' for that individual. For example, in analysis B, CHROMOPAINTER gives the counts of haplotype segments and total length of DNA for which each UK individual shares most recent common ancestry with each European individual. These values are given for chromosome 1–22 of each UK individual, and are also summed to give a genome-wide total across the autosomes (in the case of the counts data, the copying profile). Furthermore, within a UK individual, these values can be summed across any grouping of European individuals (for example those sampled from the same geographic region or assigned to the same European group, see below) providing an estimate of the counts of haplotype segments and/or total length of DNA for which each UK individual shares most recent common ancestry with any European group (a group copying profile). It is natural to average these values across UK individuals assigned to the same cluster (see below) to get average values for all UK individuals from a particular cluster; a 'copying vector' for the cluster as a whole.

For analyses A and C, described above, we used the algorithm implemented in the program fineSTRUCTURE¹⁵ to group the UK and European individuals respectively into genetically relatively homogeneous clusters. The fineSTRUCTURE program takes as its input the counts of haplotype segments for which each individual

shares recent common ancestry with every other as inferred by CHROMOPAINTER (summed across all chromosomes, the copying profile). The choice to use counts in this analysis is motivated by the underlying 'painting' model used by CHROMOPAINTER, in which segments are shared with individuals chosen independently from one another, and there is a constant switch rate between segments. Under this model, each segment provides an equal amount of independent information, while segment lengths are uninformative, so the segment counts provide a natural basis for inference, and this is why they are used. However, we note that in practice fineSTRUCTURE attempts to allow for departures from this modelling assumption (which is expected to only be an approximation) through a scaling parameter on the (log-) likelihood. Moreover we believe there is often useful information provided in, for example, the fact that segments shared between genuinely closely related groups tend to be longer on average, akin to the idea of long segments shared 'identical by descent' with respect to some founder population. Exploring and using this length information may provide an interesting topic for future work.

We initially put all of our individuals into a single cluster at iteration 0, but otherwise used default values when running fineSTRUCTURE (see ref. 15 for details). Each Markov Chain Monte Carlo (MCMC) iteration of fineSTRUCTURE provides the number of clusters and the cluster membership of each individual, sampled according to their posterior probabilities under the fineSTRUCTURE model. We sampled values every 10,000 iterations for 1 million MCMC iterations following either 1 million (analysis A) or 3 million (analysis C) 'burn-in' iterations. Starting from the MCMC sample with the highest posterior probability among all samples, fineSTRUCTURE performed 100,000 additional hill-climbing moves to reach its final inferred state.

Next we undertook an additional step to improve fineSTRUCTURE's inference for cluster membership. This is an addition to the fineSTRUCTURE algorithm¹⁵. While fineSTRUCTURE's final inferred state has been shown to give reasonable results in practice¹⁵, it relies heavily on a single MCMC sample observation. Although this single sample is the one with maximum posterior probability among all MCMC samples, the probability has been calculated assuming fixed (sampled) values for a large number of parameters that include the total number of clusters, each individual's final inferred cluster assignment, and other modelling parameters. Therefore, a concern is that the posterior distribution will be relatively flat across such an extensive state space, such that fairly divergent parameter values may result in similar posterior probabilities. In contrast, the marginal posterior distribution of each individual's cluster assignment across all MCMC runs should be substantially more informative, improving the assignment of individuals to clusters. Informally, by chance alone any given individual may not be in its own optimal (highest probability) cluster in the final inferred state, despite the overall posterior probability being at its maximum. We thus seek to reassign any such individuals to their most probable cluster. We therefore leverage the marginal information of each individual's cluster assignment from the values of the MCMC samples recorded every 10,000 iterations (see above) in order to re-assign individuals to clusters. Specifically, assuming we have N total individuals and M MCMC samples, and starting from the K clusters in fineSTRUCTURE's 'final inferred state', we performed the following procedure:

1. We find the number $x_i^{(m)}$ of individuals that cluster with individual i (including individual i itself) in MCMC sample m , for $i = 1, \dots, N$ and $m = 1, \dots, M$.
2. We furthermore find the number $y_{ik}^{(m)} \leq x_i^{(m)}$ of individuals that both cluster with individual i in MCMC sample m and that are in cluster k of the final inferred state, for $k = 1, \dots, K$.
3. We re-assign each individual i to the cluster k with the maximum value of $\sum_{m=1, \dots, M} [y_{ik}^{(m)} / x_i^{(m)}]$ across all k in $1, \dots, K$. These re-assignments give a new final inferred state; note these re-assignments can reduce the total number of clusters K .
4. We repeat steps 1–3 for 50 iterations.

This procedure gives the final cluster assignments for each individual.

One feature of this additional procedure used for reassigning individuals to clusters is that we obtain measurements of the confidence in the assignment of each individual i to each cluster k . For each individual i , the values of $\sum_{m=1, \dots, M} [y_{ik}^{(m)} / x_i^{(m)}]$ from the final iteration can be normalized across k to sum to one, and stored in the K -vector $P_{k,i}$. These quantities have a natural interpretation as a measure of the confidence associated with the assignment of individual i to each cluster k . Note that we assign individual i to the cluster k for which the value of the measurement is maximal. Call this maximal value $P_{k_{\max},i}$. It is possible to apply a threshold t , $0 < t < 1$, to the assignment of individuals to clusters so that an individual is only assigned to a cluster if $P_{k_{\max},i} > t$. If not, then the individual may be removed from subsequent analyses. We investigated the effect of setting such a threshold t . The main observation is that applying a threshold has very little effect on the make-up and distribution of clusters across the UK, nor on downstream analyses (data not shown). For further discussion see Supplementary Note.

One possible consequence of this extra procedure is to reduce the final number of clusters inferred from that of the so-called final inferred state. For analysis A,

the final number of UK clusters inferred, after the extra procedure, is 53 (the initial final inferred state had 55). For analysis C the final number of European groups inferred is 145 (no change to the initial final inferred state).

We assessed convergence of the fineSTRUCTURE MCMC runs in various ways. This included running independent chains, and comparing aspects of the assignments of individuals to clusters, and the results of downstream analyses, between the two chains. Reassuringly, given the size of the state space being explored, these diagnostics confirmed mixing of the MCMC chains (Extended Data Fig. 2).

Using the final assignments, we used fineSTRUCTURE to construct a 'tree' in the default manner described in ref. 15 by successively merging pairs of clusters. Starting at the final cluster assignments, fineSTRUCTURE merged the pair of clusters whose merging gave the smallest decrease to the posterior probability among all possible pairwise merges. This gives the next level up in the tree (with one fewer cluster). We repeat this merging process at the new level and continue until just two clusters remain. Figure 1 shows the assignment of individuals to clusters for the level of the tree when 17 clusters remain. The final cluster assignments and the assignments of individuals to clusters at all levels of the tree are provided in Supplementary Figs 1.1–1.24 for the UK clustering analyses (A). The tree so obtained is a hierarchical clustering tree and should not be interpreted as a phylogeny. Nonetheless there is information about the strength of the differentiation between clusters in these trees.

It is possible to use the vectors of measures $P_{k,i}$ defined above, of the confidence associated with the assignment of individual i to each cluster k in the final inferred state, to reassign individuals to clusters at any level of the tree. Consider the following. Define the lowest or finest level of the tree, the level relating to the final cluster assignments, to be L_K , where K is the number of clusters in the final inferred state. Then define each level of the tree to be L_j , where j in $(2, 3, \dots, K)$ is the number of clusters at the level of interest. For a given level of the tree L_j , each cluster $C_{j,i}$, j in $(1, \dots, J)$, is made up of one or more clusters at the lowest level of the tree, merged into a single cluster. For example, the large UK cluster in central and southern England at the level containing 17 clusters (depicted in Fig. 1, red squares) is the union of eleven smaller clusters from the final inferred state. For each individual i it is possible to define a new J -vector of measures $P_{j,i}$ for level L_j , where for each cluster $C_{j,i}$ we sum the values in $P_{k,i}$ for all clusters that are merged to form $C_{j,i}$, and store the result in component j of $P_{j,i}$. Thus, for our previous example of the large cluster in central and southern England at the level containing 17 clusters, for each individual i we sum the values relating to the eleven constituent clusters at the final inferred state that make up this larger cluster, and use this as the measure of confidence that the individual i is assigned to the larger cluster. We can use the vector of measures $P_{j,i}$ so-defined to reassign individuals to the cluster for which $P_{j,i}$ is maximal. This will potentially result in some individuals being reassigned to a different cluster from the one to which they were assigned by the standard tree building method. For example, we see this has occurred for exactly one individual in Extended Data Fig. 1, resulting in the different total numbers assigned to the red square and purple cross clusters in Extended Data Fig. 1 when compared to Supplementary Information Fig. 1.16 (both depicting 17 clusters). One advantage of this process is that we can interpret $P_{j,i}$ as a measure of the confidence of the assignment of an individual i to each cluster at the given level L_j . We can also set a threshold t and examine which individuals have lower confidence assignments to their cluster, where by 'lower confidence' we mean that the maximum value in the vector $P_{j,i}$ is less than t . We depict this for the UK clustering at the level of 17 clusters in Extended Data Fig. 1, when we set $t = 0.7$.

Other methods for detecting population structure. We implemented principal components analysis (PCA) using the package MMM¹⁷. We applied PCA to the intersection of the SNPs used for PCA in the WTCCC2 project²³ and the SNPs passing quality control filters in UK sample in this paper. This resulted in 188,329 SNPs with minor allele frequency >0.05 in the UK population. These SNPs are distributed approximately evenly with respect to the genetic distance across the 22 autosomes. We excluded all SNPs in regions with unusually high loadings based on visual inspection of the first 20 axes of PCA applied to the UK control samples of WTCCC2. The results are shown in Extended Data Fig. 3a.

We also applied the program ADMIXTURE¹⁶ to these same data, using default settings as recommended by the authors. The ADMIXTURE model effectively assumes independence of the markers used across the genome. We ran ADMIXTURE three times, corresponding to three different choices for the number of clusters to be used for classification (K). To understand the method in the simplest cases we set $K = 2$ and $K = 3$, and for comparison to the results presented in our main analyses we set $K = 17$. The results are shown in Extended Data Fig. 3b.

Continuous or discrete frameworks for modelling and inferring population structure. There is a general issue when modelling genetic variation from spatially structured populations as to whether to use models which characterize the population as comprised of distinct subpopulations, or at the other extreme to model

the population in continuous space, without distinct subgroups, where isolation by distance is the primary factor in giving rise to geographical substructure^{16,24,27,28}. Both are obviously oversimplifications for natural populations, and in particular for humans, and are more naturally thought of as caricatures and as endpoints of a spectrum, with debate as to which might be closer to capturing the important features of historical human demography.

One potential criticism of the fineSTRUCTURE approach is that it is embedded in a framework of discrete subgroups. There is an obvious sense in which fineSTRUCTURE is closer to this framework: it explicitly estimates a set of subgroups in the population, on the basis of patterns of shared ancestry. Although this is a description of the population, rather than a model of it, it might well be more natural or useful if there is, in reality, some underlying discreteness. On the other hand, the hierarchical tree estimated by fineSTRUCTURE allows viewing of the population at multiple levels of clustering. This does not stipulate a fixed number of subgroups, and instead provides a complex description of the underlying structure—in effect zooming in from the coarsest partition of the population as two subgroups to examine finer and finer partitions. Taken together, we argue that this approach is better suited to capturing the complexities of real populations than had it only described a single set of discrete subgroups. Our approach, of probabilistically classifying individuals into groups at a particular level, rather than forcing them to belong to exactly one cluster, also allows some flexibility in a world where there is smoother variation with geography.

Clearly some, but not all, aspects of human demography will be influenced by the dynamics of isolation by distance. Conversely, cultural, linguistic, and geographical barriers will all tend to encourage boundaries, and hence discreteness of subgroups. We are encouraged by the fact that the multi-level descriptive framework of fineSTRUCTURE, as applied to the subtle levels of population structure within the UK, is clearly capturing real effects, as evidenced for example by the concordance with geography, largely non-overlapping clusters (compared with ADMIXTURE), confident assignment of individuals to clusters in most cases (typically except where clusters overlap geographically), and its ability to detect groups which reflect known historical events.

Estimating ancestry profiles. To understand the genetic make-up of different genetic clusters in the UK with respect to potential ancestral populations we performed the following analyses. For analysis B (above) the CHROMOPAINTER algorithm provides estimates of the proportion of each UK individual's DNA that is most closely related ancestrally to each European individual, among all the sample members. These proportions can then be summed across groups. These proportions approximate the fraction of an individual's DNA that coalesces, back in time, most recently with each particular sampled individual¹⁵. Because in humans these coalescence events can be far back in time relative to population separation times, we expect them to often predate population splits (that is, we expect incomplete lineage sorting). This leads to differences in the amount of DNA copied from different European groups being subtle, in a sense adding noise. The amount of noise depends on the number of individuals sampled—and thus potentially sharing DNA—in the different groups, with larger sample sizes likely to reduce noise. In addition, we rely on informative variation patterns to identify individuals from whom DNA is copied, adding additional noise, which may systematically vary across the genome. To account for this noise we follow ref. 22, so that at each level of the hierarchical clustering tree of the UK samples, and for a fixed level (see main text) of the European samples' hierarchical clustering tree, we perform multiple linear regression as follows. For each level of the hierarchical clustering tree of the UK samples, and for the set of G ($= 51$) groups inferred for Europe we perform the following linear regression. Let Y_P be a G -vector describing the average proportion of DNA genome-wide that a cluster P of UK individuals copies from each of G groups of European individuals, as inferred by CHROMOPAINTER. That is, element g of Y_P consists of CHROMOPAINTER's total genome-wide length (in cM) of all haplotype segments inferred to be most closely related ancestrally to any individual of European group g , normalized to sum to unity across all g in $1, \dots, G$ within a UK individual and then averaged across all individuals in the UK cluster P . We use copying lengths, rather than counts (used in the clustering itself), for this analysis because all individuals have the same total genetic length, but this length may be broken into differing numbers of copying segments in different individuals. Thus it is straightforward to interpret coefficients in the below linear regression, in terms of the fraction of the genome contributed by different components in the mixture, using copying lengths, but interpretation would be more difficult using counts of shared DNA segments. Analogously, let X_g be a G -vector describing the average proportion of DNA that the European individuals of group g copy from each of the G European groups as inferred by CHROMOPAINTER, including their own group (though note individuals are not allowed to copy from their own haplotypes in CHROMOPAINTER). We assume

$$Y_P = \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_G X_G, \quad (1)$$

and solve simultaneously for the β_g under the restriction that each $\beta_g \geq 0$ and $\sum_{g=1}^G \beta_g = 1$, using a slight adaptation of the non-negative-least-squares (nnls) function in the statistical software package R (see ref. 29).

We interpret the inferred value for β_g as the average proportion of genome-wide DNA of a UK individual from cluster P that is most closely related ancestrally to European group g . We refer to these vectors as 'ancestry profiles'.

To assess statistical uncertainty in our estimates of the β_g for each UK cluster P , we perform a bootstrap procedure where we re-sample the chromosomes of the N_P UK individuals in this group (constructing pseudo-individuals by sampling pairs of chromosomes for each of the autosomes). In particular, for each bootstrap iteration, we randomly sample the G -vector of CHROMOPAINTER output across these UK individuals N_P times with replacement for each chromosome 1–22. We then generate each of N_P 'pseudo-individuals' by randomly summing 22 pairs of these samples (without replacement), one pair per chromosome, and then summing across the first, respectively the second, member of each pair before rescaling the resulting G -vectors to sum to unity. Averaging each element of the G -vectors across these N_P pseudo-individuals gives us a new re-sampled value of Y_P , which we then substitute into (1) above to generate new inferred values of the β_g . We repeat this procedure 1,000 times, reporting the inner 95% quantiles of the sampling distribution for a given European group g across these 1,000 bootstrap re-samples (see Supplementary Table 4 and Extended Data Fig. 6a).

Assessing the strength and robustness of the inferred population structure— F_{ST} , identity by descent (IBD) and total variation distance (TVD). Using the same set of SNPs that were used for the PCA analyses (see above) we analysed pairwise F_{ST} both between the sample collection districts, and between the 17 inferred clusters from our main analysis using the method implemented in the program Eigensoft³⁰. The complete matrices of pairwise F_{ST} values are given in Supplementary Tables 1 and 2.

To investigate the effect that recent shared ancestry may have on our analyses we calculated a measure of pairwise IBD and compared its distribution within clusters to its distribution across the whole sample. This measure uses a hidden Markov model (HMM) to estimate IBD across the genome¹⁴. The measure is likely to be useful when the shared relatedness is just a few generations in the past, allowing the identification of pairs of individuals in our UK sample that are reasonably closely related. The results are plotted in Extended Data Fig. 4. Reassuringly, these confirm that levels of relatedness within clusters are typically similar to those between clusters, and hence that our observed clusters are not an artefact of a sampling scheme which preferentially selected closely related individuals from regional localities.

To quantify the strength of differences between the inferred clusters we perform the following analyses. As noted above we can summarize the copying profiles of all the samples in a given cluster X to produce a characteristic 'copying vector' $x = (x_1, x_2, \dots, x_n)$; the average (across individuals in cluster X) proportion of each individual in cluster X 's closest ancestry that is attributed to individuals from each of the clusters, $Y = (Y_1, Y_2, \dots, Y_n)$, where n is the number of inferred clusters. In fact, this copying vector can be calculated for any group of samples (that is, not only the inferred clusters). One can use these vectors to test if the clusters inferred by fineSTRUCTURE are capturing significant differences in ancestry, and to give a sense of the strength of the differences observed. Given a pair of inferred clusters (A and B) and their copying vectors (a and b respectively) one can calculate the total variation distance (TVD_{CV}) between the pair:

$$TVD_{CV}(A, B) = 0.5 \times \sum_{i=1}^n (|a_i - b_i|).$$

TVD_{CV} can be interpreted as a measure of the difference between the two clusters. (As the copying vectors are discrete probability distributions over the set of clusters, total variation distance is a natural metric for quantifying the difference between them.)

Furthermore, given a pair of clusters (A and B) one can randomly reassign the individuals in the clusters, maintaining the cluster sizes, to obtain a new pair of clusters (A' and B' , of the same size as A and B , respectively). One can then calculate the copying vectors (a' and b') for the new clusters A' and B' , and the total variation distance between them. Repeating this process m times one can obtain a P value from a permutation test of the null hypothesis that, given the cluster sizes, the individuals in the two clusters are assigned randomly to each cluster. Here the P value is the proportion of the m permutations where $TVD_{CV}(A', B') \geq TVD_{CV}(A, B)$. Supplementary Table 3 shows the value of the TVD_{CV} statistic for all pairs of the 17 clusters used in our main analyses.

Similarly, rather than using the copying vectors for a pair of clusters (A and B), one can use the ancestry profiles of the clusters (α and β) to calculate the total variation distance between the ancestry profiles of a pair of clusters (TVD_{AP}):

$$\text{TVD}_{\text{AP}}(A, B) = 0.5 \times \sum_{i=1}^n (|\alpha_i - \beta_i|)$$

TVD_{AP} can be interpreted as a measure of the difference between the ancestry profiles of the two clusters. (Again, as ancestry profiles are discrete probability distributions, total variation distance is a natural metric for quantifying the difference between them.)

As above, one can permute the individuals that are assigned to each cluster, maintaining the cluster sizes, and calculate the ancestry profiles of the resulting clusters (A' and B') and the total variation distance between them. As before, repeating this process m -times one can obtain a P value from a permutation test of the null hypothesis that, given the cluster sizes, the individuals in the two clusters are assigned randomly to each cluster with respect to their ancestry profile. Here the P value is the proportion of the m permutations where $\text{TVD}_{\text{AP}}(A', B') \geq \text{TVD}_{\text{AP}}(A, B)$. Supplementary Table 5 gives TVD_{AP} for all pairs of ancestry profiles for the 17 UK clusters used in our main analyses, and gives the associated P values based on 1,000 permutations.

Assessing the accuracy and robustness of the ancestry profiles. We undertook a number of simulation studies, generating data with similar properties to the actual data, to assess the accuracy of the estimated ancestry profiles. These suggested good accuracy of the major components of our estimated ancestry profiles.

A major challenge for this kind of simulation study is in simulating data which has similar properties to the real data. The subtle similarities and differences within and between our various UK clusters and European groups are generated by their complicated shared and distinct demographic histories. This true demographic history is unknown and might not be well approximated by simple models that can be simulated from, and so it is not possible to simulate realistic data from the appropriate model³¹. Instead, we used subsamples of the real data for our simulation studies. This has the advantage that it replicates patterns in the real data, but the disadvantage that simulation studies must be based on smaller sample sizes than the actual study. (Since some of the data are needed to simulate the scenario of interest, and the rest of the data to analyse that scenario, so neither the simulated data set, nor the data used for analysis, can be as large as the actual data set.)

For each simulation scenario described below, we generated N simulated individuals as mixtures of two populations A and B intermixing λ generations ago in proportions α, β ($\alpha + \beta = 1$) respectively, closely following established approaches^{22,32,33}. Informally, to simulate an admixed haploid chromosome we did the following: a genetic distance x (in centimorgans) was sampled from an exponential distribution with rate $\lambda/100$. The first x cM of the simulated chromosome was composed of the first x cM of a real data chromosome selected randomly from either population A or B according to the proportions of admixture α and β (the specific values used are given below). Then a new genetic distance was sampled from the same exponential distribution (rate = $\lambda/100$), and the process repeated until an entire simulated chromosome was generated. This was repeated for all 22 autosomes, resulting in a single (haploid) set of chromosomes for one individual. We did this $2N$ times, generating $2N$ full sets of haploid autosomes. (To limit the chance of multiple simulated individuals copying from the same real data individual at any location in the genome, wherever possible the new piece of chromosome sampled was selected from the pool of chromosomes in the selected population (A or B) for which no other previously simulated chromosome had copied at the same location. When this was not possible, a chromosome was selected at random from the selected population (A or B), see ref. 22. Diploid individuals were constructed by aggregating two full sets of haploid chromosomes, making N simulated individuals in total.

We considered three scenarios of two populations admixing, and for each of these scenarios we considered three proportions of admixture for the second group ($\beta = 0.1, 0.25$ and 0.5). This yielded the following nine sets of simulations:

- (1) 'Italy and northern Germany': $N = 25$, $\lambda = 40$, $\beta = 0.1, 0.25$ and 0.5 , derived by mixing 30 randomly sampled individuals from the Italian Group ITA36 (which contains 284 individuals) with 10 randomly sampled individuals from GER3 (58 individuals).
- (2) 'North Wales and Norway': $N = 40$, $\lambda = 29$, $\beta = 0.1, 0.25$ and 0.5 , derived by mixing 75 individuals from the *N Wales* cluster with 10 randomly sampled individuals from NOR72 (116 individuals) and 10 from NOR71 (148 individuals).
- (3) 'North Wales and Denmark': $N = 25$, $\lambda = 40$, $\beta = 0.1, 0.25$ and 0.5 , derived by mixing 75 individuals from the *N Wales* cluster with 20 randomly sampled individuals from DEN18 (319 individuals).

These simulations were chosen both to test our model's ability to infer sources of admixture and their proportions from distinct European groups (simulation 1); as well as to mimic admixture events we infer in our main analyses, that is, relating to the Norwegian Viking (simulation 2) and Anglo-Saxon (simulation 3) migrations into the UK. Simulations (2) and (3) use samples from the *N Wales* cluster, which we infer has little evidence of DNA influx from the Norwegian Vikings and Anglo-Saxons, and mixes them with groups containing primarily individuals sampled

from Norway (2) or from Denmark (3). These simulations are used to model admixture between the 'ancient' British population (that is, genetically constituted as it was before the Saxon invasion) and Norwegian Viking or Anglo-Saxon settlers, respectively. Simulation (2) further assesses our model's ability to distinguish two distinct Norwegian sources of admixture from among 12 different groups primarily containing samples from Norway.

For each simulated data set we estimated ancestry profiles as follows. We used CHROMOPAINTER (see above) to represent each of the $2N$ simulated haplotypes as a mosaic of all the European haplotypes except those used for the relevant simulations. Specifically, the 40 samples from ITA36 and GER3 used for the simulations in the 'Italy and Northern Germany' scenario (1) were removed from the CHROMOPAINTER analysis for scenario (1). Similarly the European samples used for the simulations in (2) and (3) were removed in their respective CHROMOPAINTER analyses. This ensures that the actual admixing individuals are not sampled when forming the mosaics. We used the estimated switch and emission rates from the main analysis, described in 'Inference of population structure' above.

Recall the ancestry profiles are determined by fitting a linear mixture model that utilizes both the CHROMOPAINTER copying profiles derived from making up the 'target group' (here the simulated samples, in our main analysis the UK samples within a cluster) haplotypes from the 'source groups' (here the European samples except those used in the simulations, in our main analysis all the European samples) haplotypes, as well as the CHROMOPAINTER copying profiles used for the clustering of the 'source groups'. To obtain the latter we adapted the results from the existing CHROMOPAINTER analysis C (see 'Inference of population structure' above) as follows. (It would be computationally prohibitive to rerun the full CHROMOPAINTER analysis for each of the nine simulated data sets.) For each European individual's copying profile the elements associated with the European samples used in the simulations were removed. Then, for each of the 51 European groups, we averaged these adjusted copying profiles across all individuals assigned to the given group (excluding any individuals used in the relevant simulations) as described in 'Inference of population structure', and used the adjusted copying profiles for the 51 EU groups as covariates in our linear mixture model as described in estimating ancestry profiles.

This post-hoc adjustment of the copying profiles for each of the 51 European groups assumes that if we had repeated the CHROMOPAINTER analysis for the relevant reduced set of European samples, the copying profile of the parts of the chromosomes previously associated with the removed samples is redistributed evenly across all the other European individuals. This is inherently conservative as it is more likely that by excluding, for example, 10 of the 58 GER3 samples from the 'new' GER3 group would have resulted in an increase of copying from the other 48 GER3 samples, relative to the increase in copying from individuals from other European groups. Thus the performance of our approach for determining ancestry profiles in our simulation study is likely to be an under-representation of the performance of our approach in the main data analyses.

Furthermore, we only used a relatively small number of individuals from each of ITA36, GER3, NOR71, NOR72 and DEN18 in the simulations, to ensure a sufficient number of remaining individuals from each to use for inferring the ancestry profiles. As a consequence, the number of simulated individuals we generated is rather small, consisting of only 25 or 40 individuals per simulation, compared to our main analysis (using the real data) where many of the clusters were significantly larger. We expect the increased sample size for the majority of clusters used in our main analyses to improve our inference of ancestry profiles relative to the simulations, substantially so in some cases such as *Cent./S England* which contains 1,044 individuals.

We also adopted an alternative simulation approach for the scenarios represented by (1)–(3) above using a forwards-in-time simulation method, initialised from real data, as previously described²². In each case, we combined a subset of the same randomly sampled individuals from populations A and B above (for example, for (1), the 30 individuals from ITA36 and the 10 individuals from GER3) into a single pool population, which we then simulated forwards in time for the same λ generations as used above. To imitate the three simulations for scenario (1) above this pool population contained respectively (20, 60, 60) haplotypes from ITA36 and (20, 20, 7) haplotypes from GER3 to approximate admixture contributions of (0.5, 0.25, 0.1) from GER3. Similarly for scenario (2) the pool population contained (40, 120, 150) haplotypes from *N Wales* and (40, 40, 18) haplotypes from NOR72/NOR71 (half from each); and for scenario (3) (50, 150, 150) haplotypes from *N Wales* and (50, 50, 18) haplotypes from DEN18.

To create the next generation of haplotypes following this admixture event, we randomly sampled two distinct parental haplotypes (each comprising a full set of 22 single chromosomes from one individual) from the pool. We composed a new set of haplotypes for an individual in the next generation as a mosaic of chunks from these two parent sets, with switches in the mosaic based on the HapMap Phase 2 genetic map (June 2008, build 36 genetic map, as above). More specifically, we

determined the number of recombination breakpoints on each chromosome by summing a random sample from a Bernoulli distribution with probability 0.5 (which models the expected obligate crossover per generation per chromosome) and a random sample from a Poisson distribution with rate equal to the total genetic length of the chromosome in Morgans minus 0.5 (which models the remaining crossovers). We then sampled the physical location of each of the breakpoints independently according to their relative genetic map value, copying segments on either side of a breakpoint without mutation from the chromosome's two different parents. In the first generation after the admixture, we repeated this process to generate 500 full (that is, chromosomes 1–22) sets of haplotypes. For the remaining $\lambda - 1$ generations, 500 new full sets of haplotypes were each simulated in the same manner as a mosaic of chunks from two distinct full sets of haplotypes randomly sampled with replacement from the previous generation. After λ generations, we randomly sampled distinct haplotypes (that is, without replacement) to form N individuals for subsequent analysis, where N is the same as in the relevant scenario of (1)–(3) above. We then inferred ancestry proportions in these N simulated individuals in the same manner described above.

The resulting ancestry profiles from all 18 simulation studies (2 simulation methods times 3 scenarios times 3 admixture proportions (β)) are given in Supplementary Table 6.

Dating admixture events in Orkney and southeast England. We ran GLOBETROTTER²² to estimate the time of the major admixture events contributing to the make up of the *Cent./S England* cluster and the three clusters in Orkney (*Westray*, *Orkney 1* and *Orkney 2*) using the 51 European groups as surrogates for the putative admixing 'source groups' (that is, using analysis B from 'Inference of population structure' above) and assuming a single 'pulse' of admixture when analysing each UK cluster. We closely follow the application of GLOBETROTTER as described by the authors. In short, CHROMOPAINTER identifies the segments of DNA within each UK individual's genome that are most closely related ancestrally to each European group, as described in 'Inference of population structure'. GLOBETROTTER measures the decay of association versus genetic distance between the segments copied from a given pair of European groups. Assuming a single pulse of admixture between two or more distinct admixing source groups, theoretical considerations predict that this decay will be exponentially distributed with rate equal to the time (in generations ago) that this admixture occurred³⁴. GLOBETROTTER jointly fits an exponential distribution to the decay curves for all pairwise combinations of European groups and determines the single best fitting rate, hence determining the most likely single admixture event and estimating the date it occurred. Instead of requiring specific genetic surrogates to represent each admixing source group involved in the admixture, as in other dating approaches such as ROLLOFF³⁵, GLOBETROTTER aims to infer the haplotype composition of each source group for the admixture as a linear combination of those carried by sampled groups (that is, a linear combination of the 51 European groups). This results in the admixed group themselves automatically being represented in the same form—as a mixture of mixtures—consistent with the linear estimation procedure we applied for each UK group, before estimating admixture dates for each group.

The following provides more details on our approach for dating and estimating admixture proportions within a single UK cluster; full details of the GLOBETROTTER method are provided in ref. 22. For each haploid set of chromosomes of each individual from a given UK cluster, we consider the genome-wide mosaic inferred by CHROMOPAINTER in the UK on Europe analysis (analysis B from 'Inference of population structure' above). In this manner each UK sample's haploid genome is pieced together as a series of 'chunks', with each chunk a contiguous segment of DNA best matching a European sample inferred to be most closely related ancestrally to that segment. We note that CHROMOPAINTER infers these mosaics for each individual many times in a probabilistic manner, so we can sample from the set of mosaics for a given individual. We sampled 10 such mosaics for each haploid genome of each UK individual in the cluster we are focusing on, giving 20 total mosaics for each UK individual.

Consider two of these 20 mosaics (these two could be the same sampled mosaic). We compare each chunk on mosaic 1 to each chunk on mosaic 2. For each pair of chunks, we record the two European groups (perhaps the same) copied at each chunk (or more precisely, the group of the European individual inferred to be closest to the UK individual's chunk) in the pair and the genetic distance between the two chunks' midpoints. We remove any chunk pairs where this genetic distance is less than 1 cM (to avoid the effects of within population linkage disequilibrium confounding signals of admixture) or greater than 50 cM (as linkage disequilibrium attributable to admixture will have decayed to zero by this distance). Otherwise we round this genetic distance to the nearest 0.1 cM and assign the chunk pair a score S_{CP} equal to the product of the two chunks' sizes in centimorgans, with chunk sizes larger than 1 cM fixed to 1 cM. This scoring protocol weights chunks' contributions by their relative size, so that larger chunks contribute more to the score, but caps the contribution of any chunk to prevent inference from being dominated by

a small number of chunks. We repeat this for all chunk pairs across all $^{20}C_2 = 190$ combinations of mosaics. After doing so, for each pair of European groups, say A and B , and for each 0.1 cM bin d in $[1, 1.1, \dots, 50 \text{ cM}]$ we sum the S_{CP} values across all chunk pairs where (i) the genetic distance between the two chunks' midpoints is in d and (ii) one chunk in the pair copies A and the other copies B . We refer to this as the 'coancestry vector' for pair (A, B) , which contains one element for each d .

We repeat this tabulation for all pairs (A, B) of the 51 European groups, giving 51×51 such coancestry vectors. After a re-scaling and then a re-weighting of these coancestry vectors using the inferred ancestry profiles from 'Estimating ancestry profiles' above (that is, the β_g), this gives a set of reweighted coancestry vectors (referred to as 'observed coancestry curves' in ref. 22), that efficiently capture the decay of linkage disequilibrium attributable to admixture (see ref. 22 for details). In the course of this re-weighting, we remove European groups whose inferred ancestry contribution (β_g) to the given UK cluster is less than 0.1%, thus reducing the number of European groups remaining for consideration in our analysis. For each pair of European groups (M, N) , now a subset of the 51×51 total pairwise combinations, we label the reweighted coancestry vector v_{MN} . We fit a coancestry curve, $\Psi_{MN}(d)$, to the values in v_{MN} as follows: for each fixed pair (M, N) of European groups we fit the parametric model

$$v_{MN}(d) = \Psi_{MN}(d) + \epsilon,$$

where $\epsilon \geq 0$ is an error term and

$$\Psi_{MN}(d) = \tau_{MN} + \delta_{MN} \times e^{-\lambda d}.$$

Here λ is interpreted as the date of admixture in generations from present. GLOBETROTTER jointly estimates the values of τ_{MN} , δ_{MN} and λ that minimize the sum of the mean squared error across the curves, that is, that minimize

$$\sum_{(M,N)} \sum_d (v_{MN}(d) - \Psi_{MN}(d))^2.$$

The values of δ_{MN} carry information about which European group best represents each admixing source group (if any)—for example, positive values of δ_{MN} suggest that groups M and N often carry haplotypes representing the same true, unsampled admixing source, while negative values of δ_{MN} suggest that M and N represent different admixing sources. We use principal components and linear modelling to jointly analyse all δ_{MN} , both describing the haplotypes carried by each admixing source group as a linear combination of those carried by each of the 51 European groups, and inferring the proportion of admixture contributed from each source (see ref. 22 for details). The inferred mixing coefficients from this linear modelling, along with the inferred admixture proportions for each source, allows a new estimate of the ancestry proportions describing the given UK cluster (that is, analogous to those described in estimating ancestry profiles). We can therefore re-scale our coancestry vectors using these new ancestry proportions, giving new values of $\Psi_{MN}(d)$, from which we can re-infer the date(s) of admixture, offering improved accuracy of estimation provided the fitting procedure results in improvements in characterizing the true source groups. When analysing each UK cluster, we repeated this iterative process of ancestry proportion and date inference five times. Once these five iterations were completed, we then fixed the inferred ancestry proportions, and within each UK cluster performed 100 bootstrap re-samples of individuals' chromosomes to infer 95% confidence intervals for the actual admixture date.

Estimating the proportion of Saxon ancestry in central and southern England.

It is of interest to estimate the proportion of Saxon ancestry in our *Cent./S England* cluster. We have undertaken two separate analyses which bear on this, namely our estimated ancestry profiles and the GLOBETROTTER analysis. One challenge is that various distinct modern European groups may carry DNA which descends from the Saxons (or their ancestors), and hence be informative about the contribution of Saxon DNA to the UK.

The pattern of contributions to UK clusters from GER3, and its location in Europe in northern Germany, make it very likely to capture ancestry brought to the UK by Saxon migrants (see main text Discussion). As noted in the discussion in the Supplementary Note, some of the ancestry shared with the group DEN18 from modern Denmark could also reflect ancestry brought to the UK by the Saxon migrants. Ancestry shared with DEN18 could also have reached the UK in early migrations by land or sea, or in later migrations of the Danish Vikings. The fact that this group contributes some ancestry to all UK clusters is evidence that some of this ancestry sharing may indeed result from early migrations. The increased contribution of this group to the ancestry profiles of all the English clusters further suggests that some part also came to the UK with the Saxons.

The contribution to the ancestry of the UK clusters from FRA17, now spread throughout France, is also correlated with the contribution of GER3 and DEN18. One possible explanation for this pattern is that FRA17 also captures Saxon ancestry. Another explanation is that it represents ancestry that spread into the UK at a

different time, but into many of the same parts of the UK as the DNA from the later Saxon migrations. The Saxon migrations did not directly involve people from what is now France. There were movements of Germanic peoples, notably the Franks, into France around the time of the Saxon migration into England. The Germanic ancestry these migrations brought to what is now France would have been Frankish rather than Saxon, and it would have been diluted through mixing with the already substantial local populations. It thus seems unlikely that ancestry in the UK arising from the Saxon migrations would be better captured by FRA17 than by people now living near the homeland of the Saxons (represented by GER3)—the contribution of FRA17 is about threefold that of GER3. Further, the geographic pattern of FRA17 contributions differs from that of GER3 (which we see as very likely Saxon), in being relatively much higher in the Scottish and Orkney clusters. This is difficult to reconcile with ancestry from both groups arriving as part of the same migration event, and the substantial contribution of FRA17 in Scotland and Orkney, relative to GER3, is more likely to reflect an earlier influx into the UK, and increased time to spread geographically. Also, FRA17 did not figure as one of the source populations for the admixture event in *Cent./S England* estimated by the GLOBETROTTER analysis. We thus conclude that the contribution to the UK clusters from FRA17 is unlikely to reflect the Saxon migrations.

In the ancestry profile approach, we thus argue that the proportion of DNA in modern *Cent./S England* inherited from the Saxons is best captured by GER3 and some of DEN18, which would suggest a range of ~10% (assuming only GER3 reflected the Saxon migrations) to ~20% (assuming GER3 and all of DEN18 reflected the Saxon migrations). If we were wrong in concluding that the FRA17 contribution does not result from DNA which arrived with the Saxon migrations, so that some or all of it did reflect Saxon DNA, then the proportion of Saxon ancestry could be substantially higher (up to ~50%).

The GLOBETROTTER analysis of *Cent./S England* detected an admixture event, with a contribution of ~35% of DNA from GER3, with estimated dates for admixture somewhat after, but consistent with (see Discussion above), the known historical dates of the Saxon migrations (Extended Data Fig. 9).

There are inevitable uncertainties in both analyses due to the nature of the data – we are trying to estimate admixture proportions for events ~1,200–1,500 years ago on the basis of DNA from modern populations. Nonetheless we feel it is safe to conclude from our analyses that the proportion of Saxon ancestry in *Cent./S England* is very likely to be under 50%, and most likely in the range 10–40%.

Treatment of Eire. We explicitly excluded samples from Eire (the Republic of Ireland) from our European analyses, and as possible contributors to the ancestry profiles of the UK clusters, principally to allow assessment of the major migrations from continental Europe into the UK. Detailed early analyses, which included samples from Eire with the other European samples, provided evidence of shared Irish ancestry with our UK samples, presumably reflecting in part migrations from Great Britain into Eire and vice versa. Eire thus acts as a source and a sink for ancestry from the UK, which severely complicates interpretation of estimated ancestry profiles, since sharing of ancestry with Eire could reflect British migration into Eire rather than the converse. Also, the UK and Eire could share ancestry because both descend from some similar ancestral populations.

While there is historical evidence of migration after the collapse of Roman rule from Devon and Cornwall into what is now Brittany in north-west France, this does not leave a signal in, and hence does not confound, our ancestry analyses, either because we do not have appropriate samples from Brittany or because the amount of DNA transferred from Britain to France via this route is relatively small. (Had this been an effect we would have expected to see either or both of our Devon and Cornwall clusters sharing substantially more ancestry from one of the groups in France, but this was not the case.)

Maps and visualization. For the UK map boundaries we used a map of the UK sourced from the Office for National Statistics (England and Wales); National Records of Scotland; and the Northern Ireland Statistics and Research Agency. The European maps were sourced from Eurostat. For context we added the boundaries of the Republic of Ireland and the Isle of Man to the UK maps, taken from the European maps. Map boundaries were obtained in digitised form^{36–39} and were drawn using various packages in the statistical software language R.

The latitude and longitude for each UK sample's grandparents' birthplaces was assigned (geocoded) automatically¹³ using a place name gazetteer from Edina (<http://www.edina.ac.uk>). All locations were checked for consistency between project records and the automatic geocoding, and any discrepancy resolved in favour of the project records. For the UK cluster analyses shown in Fig. 1, each sample was assigned, and plotted at, the average of the latitudes and longitudes of its grandparents' birthplaces. For clarity of display a small, random, amount of noise was added to point's latitude and longitude to avoid over-plotting. Independently across points, a random value was drawn from a uniform distribution on $(-20a, 20a)$, where a is the smallest non-zero difference in latitude observed between the locations of any pair of points, and the resulting value was then added to the latitude of

the point. An analogous procedure was then performed, independently, for the longitude of each point.

For the tree depicting the order of the hierarchical merging of clusters in Fig. 1, the lengths of the branches relate to changes in the posterior of the fineSTRUCTURE model. They do not relate directly to time or other measures of genetic distance so caution is needed in their interpretation. Some additional length was added to the tips of the tree for clarity.

The ellipses displayed in Fig. 1, Extended Data Figs 1, 3 and 4 and Supplementary Fig. 1 were obtained by fitting a two dimensional t -distribution with five degrees of freedom to the plotted spatial locations associated with each cluster. Each ellipse depicts the 90% probability region of the fitted distribution.

Only limited geographic information was known about the European samples: often this was just the city or region from which the samples were taken, but sometimes only a country was known. To visualize the spatial patterns of the European genetic groups obtained from the fineSTRUCTURE analyses we plotted the European samples on a map of Europe, with colours reflecting the groups assigned by fineSTRUCTURE. We did this in two ways, plotting individual points for Fig. 2 (depicting the ancestry profiles) and using pie charts in Extended Data Fig. 5b.

For Fig. 2 we restricted ourselves to plotting only those European samples that have some fine-scale location information (that is, more precise than just country of sampling), as these samples will be informative for assessing regional fine structure (although all samples in the group are used for generating the ancestry profile). As all the samples from a given region/city have exactly the same location assigned to them, we added some random noise to each sample's assigned latitude and longitude to enable visualization on a map. To do this for each sample we drew two samples from a uniform distribution on $(-0.5, 0.5)$, in units degrees of latitude and longitude, and added the results to the sample's latitude and longitude respectively. We plot each sample as a point on the map, coloured to indicate the European group to which they are assigned. Figure 2 shows the locations of the samples assigned to each European group that contribute at least 2.5% to at least one of the UK clusters. As several inferred European groups are represented in the French sampling locations, and would thus be difficult to discern, the points for groups FRA12, FRA14 and FRA17 have been shifted by one degree of both latitude and longitude (for FRA12, -1 degree of longitude and -1 degree of latitude; FRA14, $+1$ and $+1$; FRA17, -1 and $+1$). In Fig. 2 the lines to each group (or set of groups) end at the centre of mass of the groups. This was calculated before any samples had their locations shifted (as for the French groups, and/or by adding random noise). For the Norwegian groups and the Swedish groups the line ends at the average position of the centres of mass of the constituent groups. For the groups GER3 and GER6 the centre of mass is calculated using only those samples from Germany. This is because several samples from these groups are assigned to Stockholm, Copenhagen and Oslo, all of which are major cities. We assume these samples are migrants from Germany, and thus including them would skew the centre of mass position that we interpret as the approximate historical locus for the group. This potential problem caused less of an issue for the other groups depicted in Fig. 2.

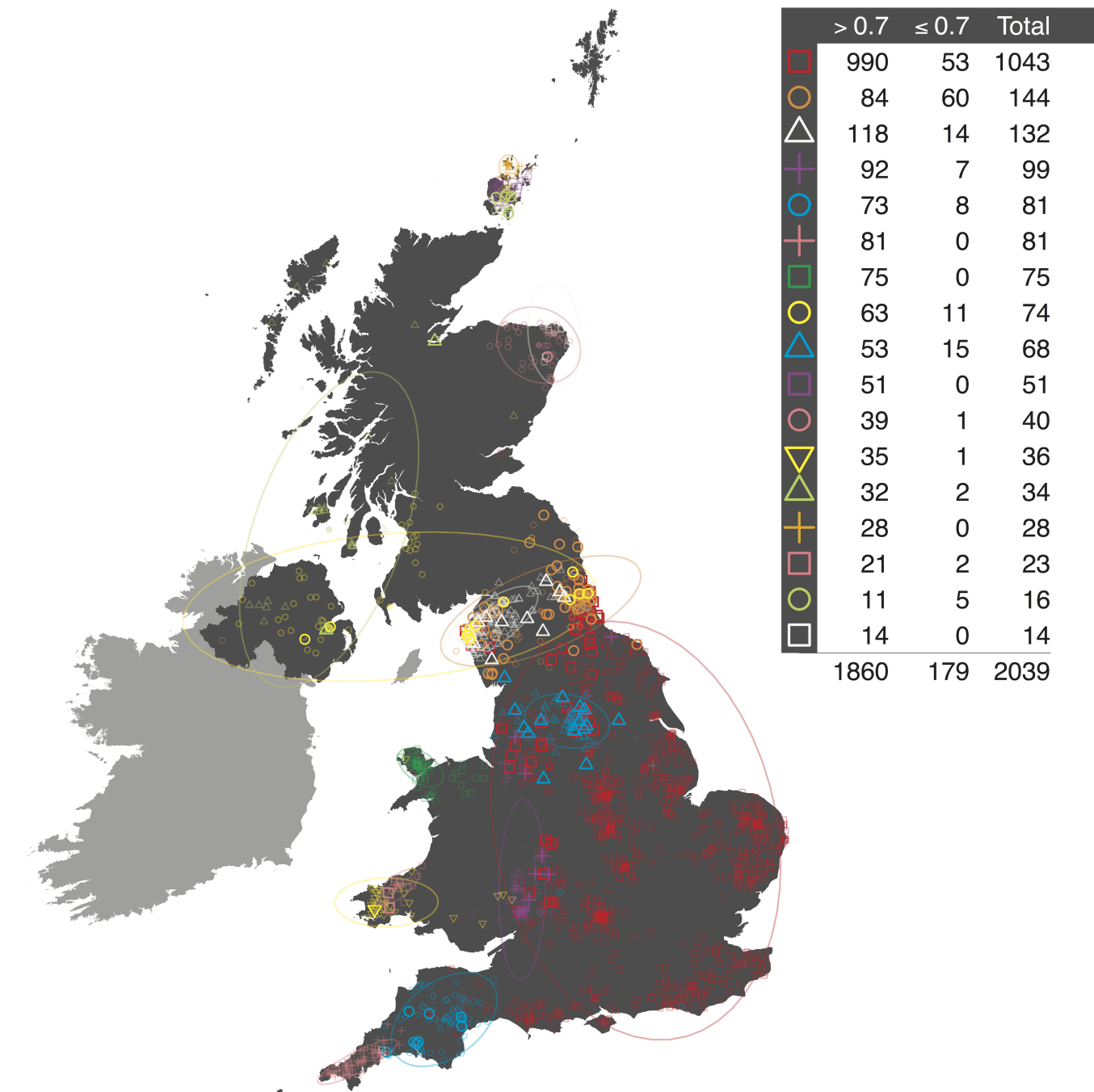
For Extended Data Fig. 5b, the spatial patterns of the European genetic groups obtained from the fineSTRUCTURE analyses are displayed in pie charts. All of the samples from the same location are displayed together in a single pie chart, with the sectors of the pie chart coloured to reflect the proportion of samples from that location that are assigned by fineSTRUCTURE to a given group. The pie charts are centred at suitable locations on the map of Europe, depending on the geographic information known (see relevant figure captions). The size of the pie chart indicates the number of samples represented by the chart. The number of samples represented by a chart is proportional to the area of the chart. For the larger sampling locations, if a European group accounts for at least 20% of a location's samples then the European group number is also displayed on the edge of the appropriate sector of the associated pie chart.

For the ancestry profile analyses we also display pie charts, this time on a map of the UK (Extended Data Fig. 6). Here each pie chart relates to one of the inferred UK clusters and is displayed at the centre of the cluster's associated ellipse (as described above). Each sector of the pie chart is coloured (and sometimes numbered) by the relevant colour (and number) of the European group it relates to. The subtended angle of each sector represents the proportion of the UK cluster's ancestry that is most similar to that of samples from the relevant European group as described in the estimating ancestry profiles section above.

Consent and study protocol. Informed consent was obtained from all subjects. For the UK subjects, ethics approval was granted by the NRES Committee, Yorkshire and the Humber – Leeds West, UK (Reference 05/Q1205/35) in March 2005. For the European subjects, informed consent and ethics approval was obtained as part of the WTCCC2 multiple sclerosis study¹⁴.

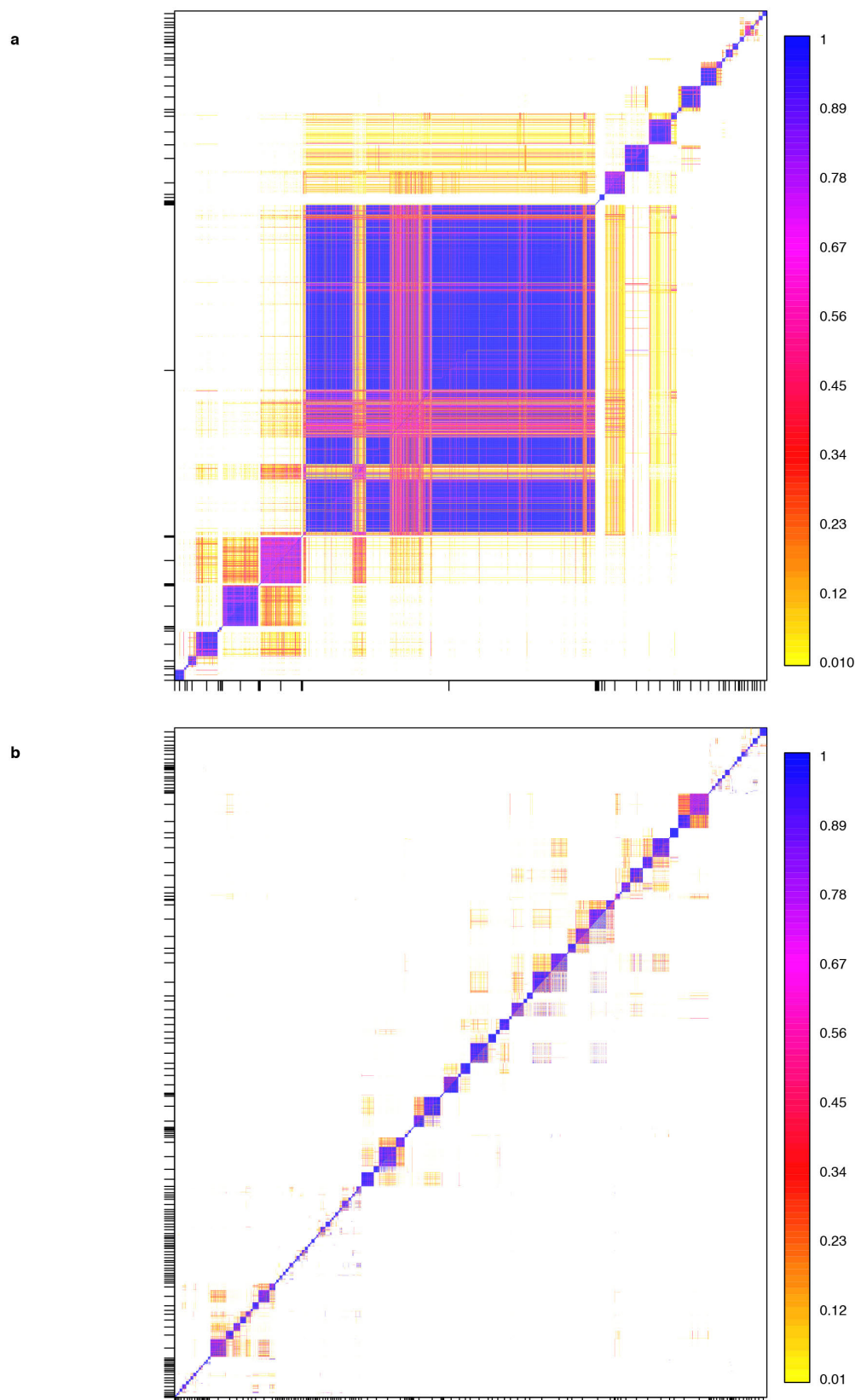
Code availability. The software CHROMOPAINTER, fineSTRUCTURE and GLOBETROTTER are available for download at (<http://www.paintmychromosomes.com/>).

23. Genetic Analysis of Psoriasis Consortium & the Wellcome Trust Case Control Consortium 2. A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. *Nature Genet.* **42**, 985–990 (2010).
24. Pritchard, J. K., Stephens, M. & Donnelly, P. Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
25. Bodmer, J. G. in *Population Structure and Genetic Disorders* (eds Eriksson, A. W., Forsius, H., Nevanlinna, H. R., Workman, P. L. & Norio, R. K.) 211–238 (Academic Press, 1980).
26. Howie, B. N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
27. Wright, S. Isolation by distance. *Genetics* **28**, 114–138 (1943).
28. Rousset, F. *Genetic Structure and Selection in Subdivided Populations* (Princeton University Press, 2004).
29. Lawson, C. L. & Hanson, R. J. *Solving Least Squares Problems* (Reprinted by the Society for Industrial and Applied Mathematics, 1995).
30. Bhatia, G., Patterson, N., Sankararaman, S. & Price, A. L. Estimating and interpreting FST: the impact of rare variants. *Genome Res.* **23**, 1514–1521 (2013).
31. Pritchard, J. K. & Przeworski, M. Linkage disequilibrium in humans: models and data. *Am. J. Hum. Genet.* **69**, 1–14 (2001).
32. Price, A. L. *et al.* Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet.* **4**, e236 (2008).
33. Moorjani, P. *et al.* The history of African gene flow into Southern Europeans, Levantines, and Jews. *PLoS Genet.* **7**, e1001373 (2011).
34. Falush, D., Stephens, M. & Pritchard, J. K. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567–1587 (2003).
35. Patterson, N. *et al.* Ancient admixture in human history. *Genetics* **192**, 1065–1093 (2012).
36. National Records of Scotland. 2011 Census: Digitised Boundary Data (Scotland) [computer file]. UK Data Service Census Support. Downloaded from: <http://edina.ac.uk/census>.
37. Northern Ireland Statistics and Research Agency. 2011 Census: Digitised Boundary Data (Northern Ireland) [computer file]. UK Data Service Census Support. Downloaded from: <http://edina.ac.uk/census>.
38. Office for National Statistics. 2011 Census: Digitised Boundary Data (England and Wales) [computer file]. UK Data Service Census Support. Downloaded from: <http://edina.ac.uk/census>.
39. European maps. Sourced from Eurostat, copyright EuroGeographics for the administrative boundaries <http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units>.



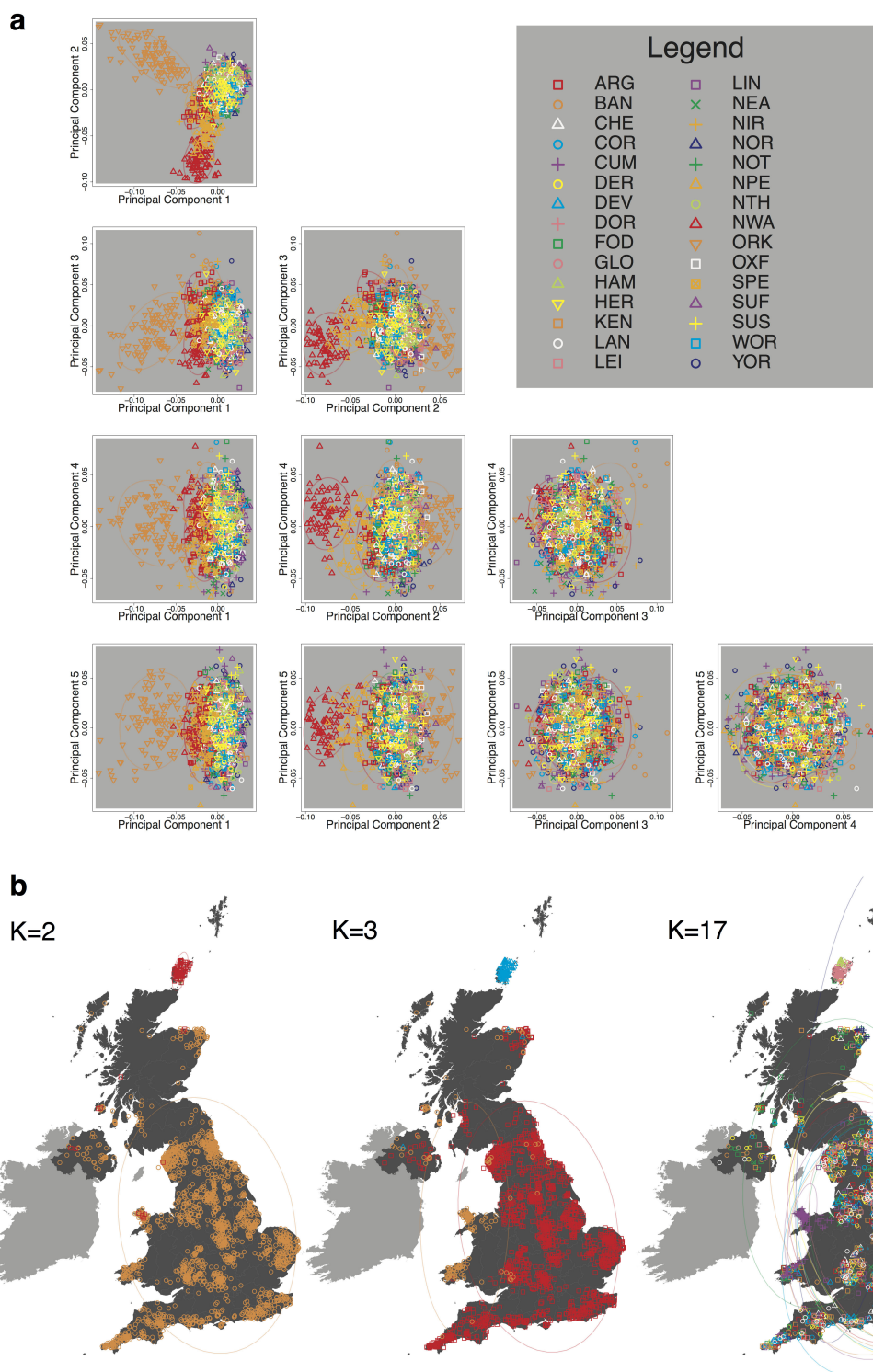
Extended Data Figure 1 | The effect of setting a threshold on the confidence of cluster assignment for the genetic clusters in the UK inferred by the fineSTRUCTURE analysis. The UK map depicts the clustering of the 2,039 UK individuals into 17 clusters on the basis of genetics alone. See Fig. 1 for further details. Here a threshold is set on the measurement of confidence used for assigning individuals to clusters (see Methods). This measure is defined on the interval $[0, 1]$, where the value 1 is interpreted as meaning complete certainty of cluster assignment and 0 as being complete lack of certainty. The plot illustrates the effect of setting a threshold of 0.7 so that a UK individual is only assigned to a cluster if the measure of assignment for that individual is greater than 0.7. All of the samples that have small, faded symbols are assigned

to their clusters with confidence greater than 0.7. Those samples for which the assignment is less confident (that is, the measure is less than or equal to 0.7) are plotted with large, bold symbols. The table shows the number of individuals with confidence measure above and below the 0.7 threshold together with the total for each UK cluster. The slight discrepancy between the totals in this figure and Supplementary Information Fig. 1.16 is due to differences in the method for assigning individuals to clusters (see Methods). The threshold of 0.7 was chosen for illustrative purposes only. Similar patterns relate to other thresholds. Contains OS data © Crown copyright and database right 2012. © EuroGeographics for some administrative boundaries.



Extended Data Figure 2 | Convergence of the algorithm implemented in fineSTRUCTURE. The fineSTRUCTURE clustering algorithm was run twice on the UK samples (a) and twice on the European samples (b) to assess convergence. The displayed heatmap depicts the proportion of sampled MCMC iterations for which each pair of UK individuals is assigned to the same cluster. The values above and below the diagonal represent two different runs of fineSTRUCTURE. Individuals are ordered along each axis according to the

inferred tree from the fineSTRUCTURE run above the diagonal, with tick marks on the axes at the middle of each cluster. Comparison between runs is made by comparing the plot above the diagonal (run two) with that below the diagonal (run one). The high degree of symmetry in the plot confirms the similarity between the runs and hence that each MCMC run has converged to very similar clusters.



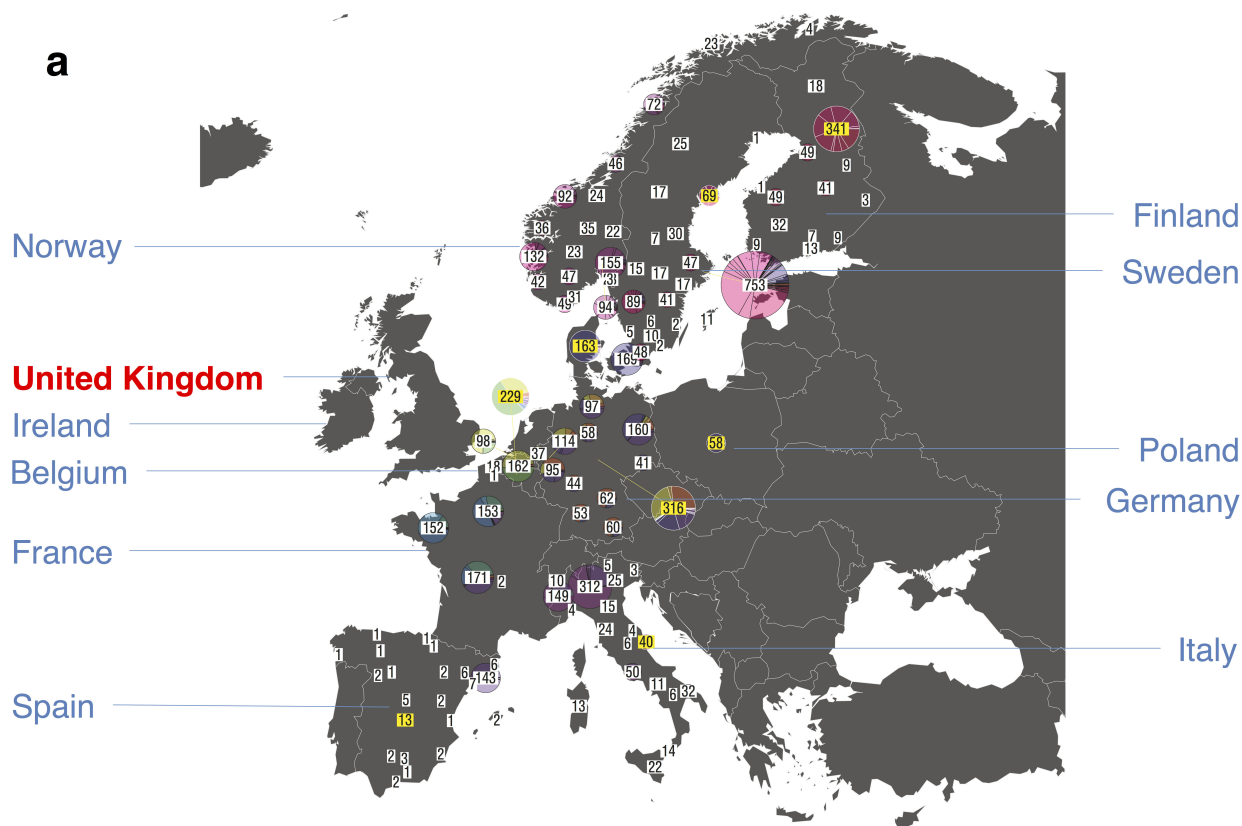
Extended Data Figure 3 | Application of standard methods for detecting population structure to the UK data. **a**, Genome-wide principal component analysis of the UK samples. The UK samples plotted against all pairs of principal component axes, for the first five axes, as determined in the genome-wide principal components analysis. Each individual is depicted by a symbol representing the district from which it was collected. The labels of the sample collection districts are interpreted as follows: CUM, Cumbria; LIN, Lincolnshire; NEA, north east England; OXF, Oxfordshire; YOR, Yorkshire; CHE, Cheshire; NTH, Northamptonshire; NOT, Nottinghamshire; DOR, Dorset; SUS, Sussex; NOR, Norfolk; WOR, Worcestershire; DEV, Devon; SPE, south Pembrokeshire; COR, Cornwall; NWA, north Wales; ARG, Argyle and Bute; NPE, North Pembrokeshire; BAN, Banff and Buchan; NIR, Northern Ireland; ORK, Orkney; SUF, Suffolk; LEI, Leicestershire; FOD, Forest of Dean; HER, Herefordshire; HAM, Hampshire; DER, Derbyshire; LAN, Lancashire;

KEN, Kent; GLO, Gloucestershire. **b**, Clustering the UK samples using the program ADMIXTURE. ADMIXTURE was applied in three scenarios, corresponding to different preset values for K , the number of clusters into which the UK samples are divided. Here $K = 2, 3$ and 17 (see Methods). A map is shown for each value of K . Each symbol on the map corresponds to one of the sampled individuals and is plotted at the centroid of their grandparents' birthplace. Each cluster is represented by a unique combination of colour and plotting symbol, with individuals depicted with the symbol of the cluster to which they are assigned. The ellipses centred on each cluster give a sense of the extent of the cluster by showing the 90% probability region of the two-dimensional t -distribution (5 degrees of freedom) which best fits the locations of the individuals in the cluster. Contains OS data © Crown copyright and database right 2012. © EuroGeographics for some administrative boundaries.

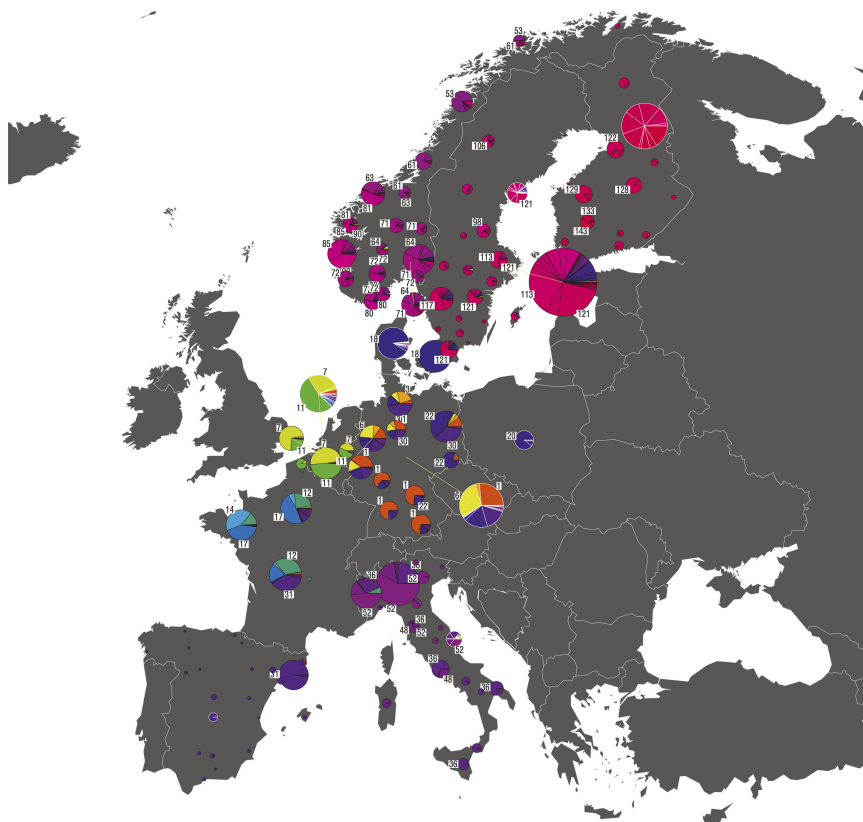


Extended Data Figure 4 | Potential recent shared ancestry in the genetic clusters in the UK inferred by the fineSTRUCTURE analysis. a. The UK map on the left depicts the clustering of the 2,039 UK individuals into 17 clusters on the basis of genetics alone. See Fig. 1 for further details. Pairwise identity by descent (IBD) within clusters and across the whole UK sample for all of the 2,039 UK individuals is shown to the right. For each of the inferred UK clusters a box and whisker plot shows the distribution of the pairwise IBD statistic (see Methods). Each box is filled by the colour of the cluster to which it relates, and the outlier points have the same shape as the cluster to which they relate. For comparison the distribution of the pairwise IBD statistic across the whole

UK sample is shown on the far right, with the box coloured grey. The light grey horizontal lines indicate the upper and lower quartiles of the IBD statistic's distribution for the whole UK sample. Along the x axis the number of samples in the associated cluster is shown. The y axis gives the value of the pairwise IBD statistic. **b.** The same information as **a** but with 53 clusters of UK individuals. Note that only clusters of size 4 or less depart substantially from the average relatedness. Contains OS data © Crown copyright and database right 2012. © EuroGeographics for some administrative boundaries.



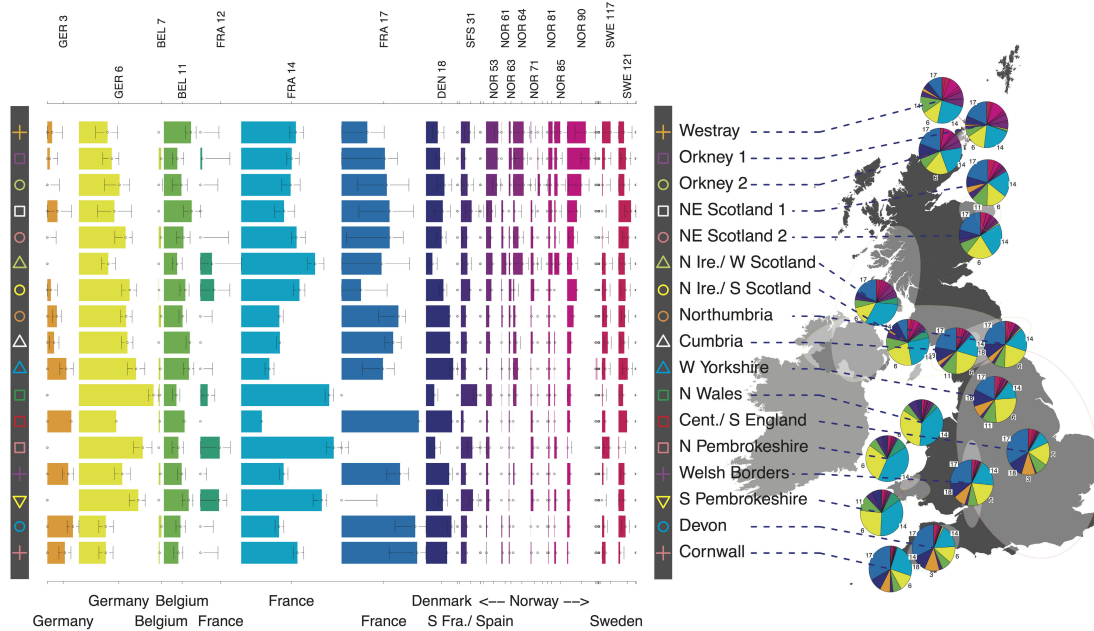
b



Extended Data Figure 5 | Population structure in the European samples.

a, Number of samples derived from each European sampling region. The 6,209 European samples used for the analyses were sampled from ten countries and various locations within each country. Each sample has a specific sampling location (often a city, but in some cases only a whole country). The numbers shown give the number of samples derived from a particular location. Some numbers are depicted out of position for clarity. In these cases a line leads from the number to the actual location. Where the sample locations are well-localized (for example, the city of sampling is known) the box surrounding the number is white. When only information about the country of sampling is known the box is coloured yellow. The numbers are overlain on a faded version of the pie charts from panel **b** for easy reference. **b**, European population structure inferred by fineSTRUCTURE. The 6,209 European samples divided into 51 genetic groups (represented by colours and labelled with a subset of the numbers between 1 and 145) using fineSTRUCTURE. For

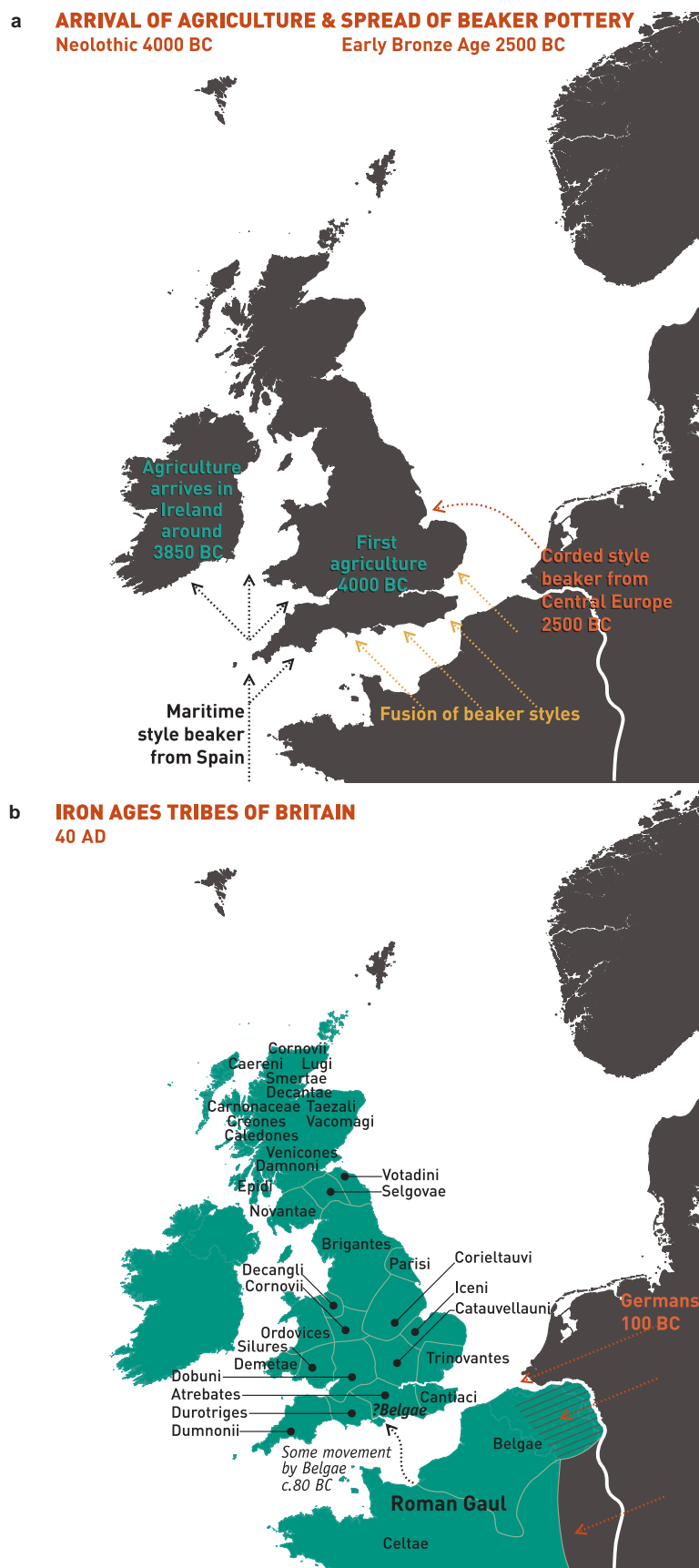
clarity the colour space has been skewed to emphasize the differences between groups 1 to 18 as these groups are the major contributors to the ancestry profiles of the UK clusters. Each sample has a specific sampling location (often a city, but in some cases only a country, see panel **a**). The pie charts are located at these sampling locations, and depict the proportion of the samples from that location assigned to each of the 51 genetic groups. Each genetic group also has a label number, which is displayed for the larger sectors of each of the pie charts. The area of the pie chart is proportional to the number of samples from that location. Pie charts with black borders correspond to well-localized samples. In contrast, for samples where only the country of sampling is known, they are combined in a single pie chart for the country, which is shown with white borders. Some pie charts are depicted out of position for clarity; in these cases a line leads from the chart to the actual location. © EuroGeographics for the administrative boundaries.

a**b**

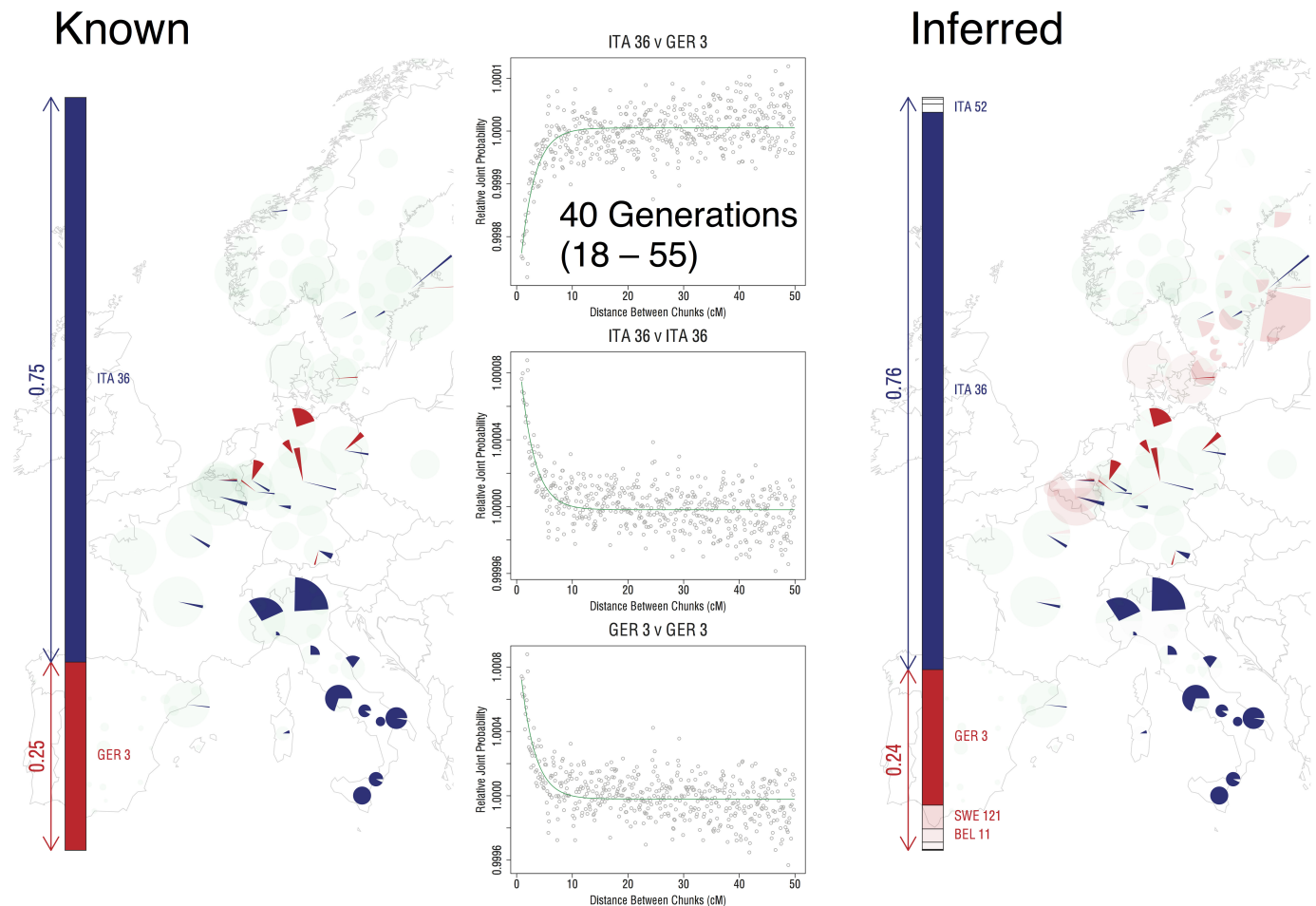
Extended Data Figure 6 | European ancestry profiles of the UK clusters.

a. The map of the UK shown relates to the map with 17 UK clusters shown in Fig. 1. Ellipses indicate the extent of the UK clusters as in Fig. 1. The pie charts represent the ancestry profile of the UK clusters from Fig. 1. Each pie chart is plotted at the centroid of the corresponding cluster, although some pie charts have been moved for clarity; in the cases where the relocation is substantial a red line leads from the pie chart to the centroid. The sectors of the pie charts are coloured with the colours of the European genetic groups (for the larger sectors the number of the European group is also given). They indicate the ancestry profiles of each UK cluster, namely the proportion of the cluster ancestry that is best represented by each of the European groups. The magnitude of the angle of a sector is proportional to the contribution of that European group to the ancestry profile of the associated UK cluster. The symbols in the grey bar to the left of the map represent the UK clusters as in Fig. 1. The bar chart in the left part of the plot depicts the same ancestry profiles of the UK clusters in a different way. Each row represents a UK cluster (arranged roughly north to south) with the symbols for the clusters from Fig. 1 indicated at each end of the

row. Each column represents a European group, with group numbers listed with a three letter prefix that, for clarity, relates to the country or countries where the cluster is most represented. The colour of each bar also indicates the European group to which the bar relates. Confidence intervals (95%) obtained from 1,000 bootstraps of the ancestry profile analysis (see Methods) are indicated on each bar. **b.** Renormalized ancestry profiles of the UK clusters illustrating possible early European contributions to the UK population. A representation of the relative contributions to the UK clusters from the three European groups (GER6-W. Germany, BEL11- Belgium, and FRA14-NW France) hypothesized to be the major contributors to the earliest migrations into the UK after the last ice age from which DNA survives to the present in substantial proportions (see Supplementary Note). Interpretation of the map, pie charts and bar chart is as for **a**. In this case, however, the proportions were renormalized to sum to 1 for the contributions from GER6, BEL11 and FRA14. Contains OS data © Crown copyright and database right 2012. © EuroGeographics for some administrative boundaries.



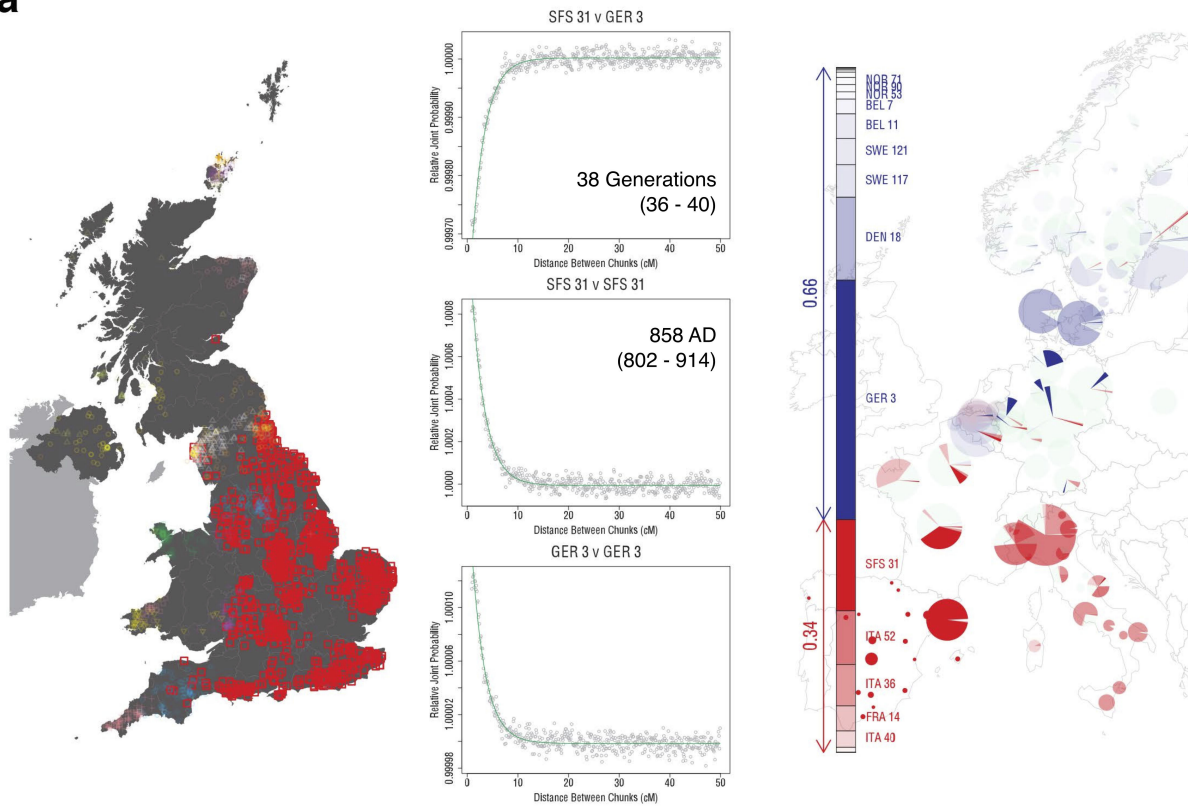
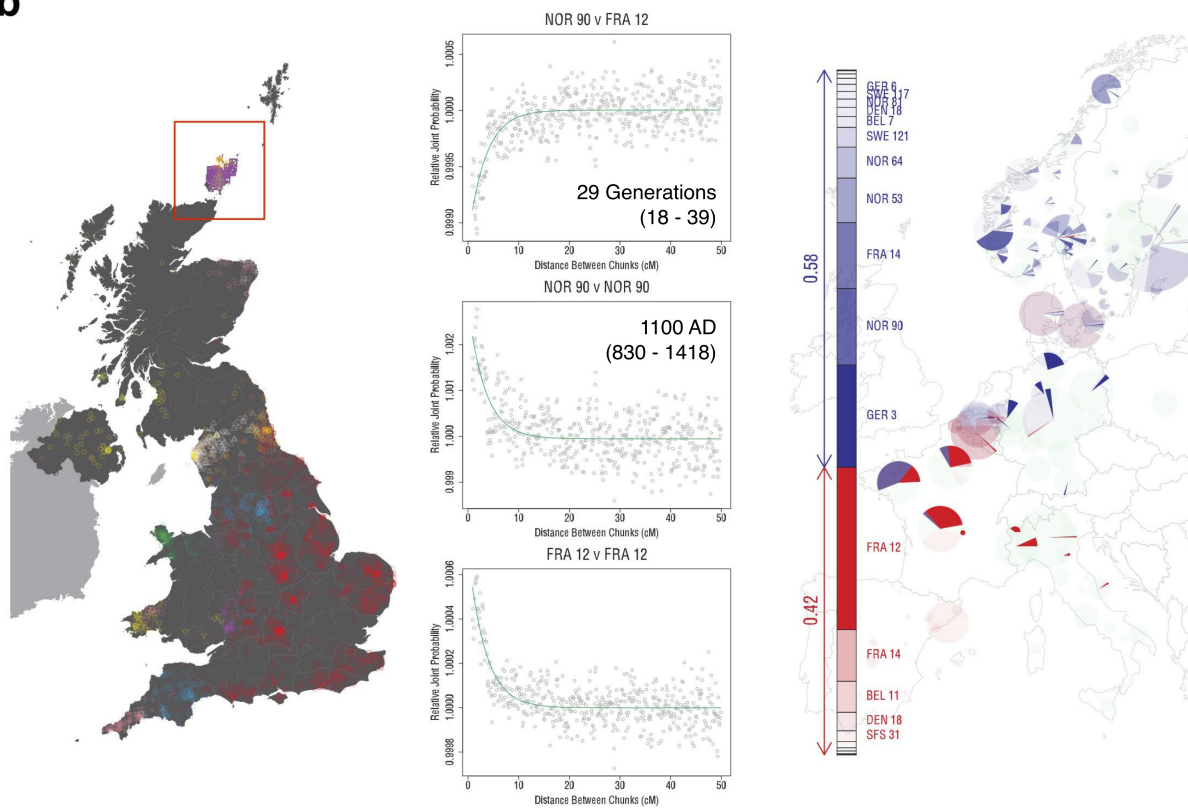
Extended Data Figure 7 | More major events in the peopling of the British Isles. See Supplementary Note for further details. **a**, The arrival of agriculture and subsequent migrations from 4000–2500 BC. **b**, The major iron age tribes of Britain around the year 40. © EuroGeographics for the administrative boundaries (coastlines).



Extended Data Figure 8 | Application of GLOBETROTTER to infer simulation of ancestry 40 generations ago between groups from Northern Germany (GER3, 25%) and Italy (ITA36, 75%). Twenty-five admixed individuals were simulated, and the individuals used to construct these simulated individuals were then removed from the list of potential donors (see Methods). Left barplot, and map: the barplot shows the true population and proportion contributed for each of the two admixing groups. The map shows, for each of the European sampling locations, the true proportion of individuals sampled from that location assigned to each of the admixing groups, coloured according to the barplot. Central three plots: example curves constructed by GLOBETROTTER to infer admixture times, and infer details of admixing groups (see Methods and Supplementary Note). For each pair of populations A and B (A can be the same as or different from B) the points show the empirical probability, relative to under independence, as a function of genetic distance x , that two positions separated by distance x correspond to ancestry donated by population A, and by population B, respectively. The green line shows GLOBETROTTER fitted exponential decay curves for the underlying (that is, expected) value of this relative probability estimate. Under a model of a single admixture event occurring g generations ago, this probability decays at a rate g according to theory, providing an estimate of the admixture time (and 95% CI) shown overlaying the curve ITA36 versus GER3. If ancestries A and B associate with the same admixing group, for example, whenever $A = B$ the fitted curve will have a negative slope, as seen for

the GER3 versus GER3 plot. If a positive slope is seen, as for the ITA36 versus GER3 plot, this implies these populations contribute to the two different respective admixing groups. Right bar-plot, and map: GLOBETROTTER produces an inference of the genetic composition of (haplotypes carried by) the two admixing groups, as a mixture of (haplotypes carried by) populations actually sampled. This mixture inference jointly uses curves for pairs of sampled populations, and the overall haplotypic makeup of different sampled populations, including the admixed group. The bar-plot shows the inferred mixture representation (dominated in each case by the true admixing groups) and estimated admixture proportion (24%, close to the truth of 25%), with more red/blue populations respectively giving a larger contribution. The map shows populations inferred as contributing to the first (pink/red shades) or second (blue shades) admixing group, respectively, as for the left map, with populations coloured according to the bar-plot. This shows populations falsely inferred as contributing material to the admixing groups were still sampled, mainly, from locations close to those of the true admixing groups. We caution that in this setting of admixture between genetically similar European groups, estimation of admixture fraction is very uncertain (see Methods) (for example, contributing populations are often impossible to definitively assign to a 'side' of the event). For further details of the analysis, for example, tests for admixture presence in this simulation, see Methods and Supplementary Note. © EuroGeographics for the administrative boundaries.

a

**b**

Extended Data Figure 9 | Application of GLOBETROTTER to infer details of admixture in the UK clusters. **a**, Inferring admixture in a population of 1,044 UK individuals from central and southern England. Left hand plot: the bold red squares show mean grandparental birthplace for each individual in this cluster. Central three plots: example curves constructed by GLOBETROTTER to infer admixture times, and infer details of admixing groups (see Methods and Supplementary Note). For each pair of populations A and B (A can be the same as or different from B) the points show the empirical probability, relative to under independence, as a function of genetic distance x , that two positions separated by distance x correspond to ancestry donated by population A, and by population B, respectively. The green line shows GLOBETROTTER fitted exponential decay curves for the underlying (that is, expected) value of this relative probability estimate. Under a model of a single admixture event occurring g generations ago, this probability decays at a rate g according to theory, providing an estimate of the admixture time (and 95% CI) shown overlaying curves SFS31 versus GER3 and SFS31 versus SFS31. If ancestries A and B associate with the same admixing group, for example, whenever $A = B$ the fitted curve will have negative slope, as seen for the GER3 versus GER3 plot. If a positive slope is seen, as for the SFS31 versus GER3 plot, this implies these populations contribute to the two different respective admixing groups. Right bar-plot, and map: GLOBETROTTER inference shows one possibility for the genetic composition of (haplotypes carried by) the two unsampled historical admixing groups, as a mixture of (haplotypes carried by) populations actually sampled. This mixture inference jointly uses curves for pairs of sampled populations, and the overall haplotypic makeup of different sampled populations, including the admixed group. The bar-plot shows the inferred mixture representation (with largest contributions in each case by GER3/DEN18, sampled most frequently from northern Germany and Denmark, and SFS31/ITA52, sampled mainly from southern France and Spain and northern Italy) and estimated admixture proportion (34%), with more intense red/blue populations respectively implying a larger contribution. The map shows populations inferred as contributing to the first (pink/red shades) or second (blue shades) admixing group respectively, with populations coloured according to the bar-plot. We caution that in this setting of admixture between genetically similar European groups, estimation of admixture fraction is very uncertain (see Methods and Supplementary Note) (for example, contributing populations are often impossible to definitively assign to a side of the event), so that other closely related scenarios, for example, a somewhat lower admixture fraction from a more completely 'GER3'-like group than that inferred, are likely consistent with the GLOBETROTTER results seen. **b**, Inferring admixture in a population of 51 UK individuals from Orkney. Left hand plot: the bold purple squares show mean grandparental

birthplace for each individual in this cluster. Central three plots: example curves constructed by GLOBETROTTER to infer admixture times, and infer details of admixing groups (see Methods and Supplementary Note). For each pair of populations A and B (A can be the same as or different from B) the points show the empirical probability, relative to under independence, as a function of genetic distance x , that two positions separated by distance x correspond to ancestry donated by population A, and by population B, respectively. The green line shows GLOBETROTTER fitted exponential decay curves for the underlying (that is, expected) value of this relative probability estimate. Under a model of a single admixture event occurring g generations ago, this probability decays at a rate g according to theory, providing an estimate of the admixture time (and 95% CI) shown overlaying curves NOR90 versus FRA12 and NOR90 versus NOR90. If ancestries A and B associate with the same admixing group, for example, whenever $A = B$ the fitted curve will have negative slope, as seen for the NOR90 versus NOR90 plot. If a positive slope is seen, as for the NOR90 versus FRA12 plot, this implies these populations contribute to the two different respective admixing groups. Right bar-plot, and map: GLOBETROTTER inference shows one possibility for the genetic composition of (haplotypes carried by) the two unsampled historical admixing groups, as a mixture of (haplotypes carried by) populations actually sampled. This mixture inference jointly uses curves for pairs of sampled populations, and the overall haplotypic makeup of different sampled populations, including the admixed group. The bar-plot shows the inferred mixture representation (with largest contribution in each case by GER3/NOR90, sampled most frequently from northern Germany and Norway, and FRA12/FRA14, both sampled mainly from France) and estimated admixture proportion (42%), with more intense red/blue populations respectively implying a larger contribution. The map shows populations inferred as contributing to the first (pink/red shades) or second (blue shades) admixing group respectively, with populations coloured according to the bar-plot. We caution that in this setting of admixture between genetically similar European groups, estimation of admixture fraction is very uncertain (see Methods and Supplementary Note) (for example, contributing populations are often impossible to definitively assign to a 'side' of the event). In particular, inspection of curves involving GER3 does not yield a clear 'side' of the event for this population, unlike the NOR90 versus FRA12 case that implies French-like and Norwegian-like haplotype presence must occur mainly in distinct admixing groups. Therefore the GER3 component might in fact capture haplotypes for either (or both) the French-like or Norwegian-like admixing groups, and the inferred scenario shows only one possibility. Contains OS data © Crown copyright and database right 2012. © EuroGeographics for some administrative boundaries.

Visualizing transient Watson–Crick–like mispairs in DNA and RNA duplexes

Isaac J. Kimsey¹, Katja Petzold², Bharathwaj Sathyamoorthy¹, Zachary W. Stein³ & Hashim M. Al-Hashimi¹

Rare tautomeric and anionic nucleobases are believed to have fundamental biological roles, but their prevalence and functional importance has remained elusive because they exist transiently, in low abundance, and involve subtle movements of protons that are difficult to visualize. Using NMR relaxation dispersion, we show here that wobble dG•dT and rG•rU mispairs in DNA and RNA duplexes exist in dynamic equilibrium with short-lived, low-populated Watson–Crick–like mispairs that are stabilized by rare enolic or anionic bases. These mispairs can evade Watson–Crick fidelity checkpoints and form with probabilities (10^{-3} to 10^{-5}) that strongly imply a universal role in replication and translation errors. Our results indicate that rare tautomeric and anionic bases are widespread in nucleic acids, expanding their structural and functional complexity beyond that attainable with canonical bases.

Nucleic acid bases exist predominantly in one neutral tautomeric form. This in turn gives rise to the strict Watson–Crick (WC) pairing rules (Fig. 1a) that govern how genetic information is replicated, transcribed, and translated. However, if bases adopt alternative energetically disfavoured tautomeric or anionic forms (Fig. 1a), pairing rules can be violated and new functions can emerge. For example, although rarely observed, minor tautomeric and anionic bases can form WC-like dG•dT^{1–3}, dA•dC⁴, and rG•rU^{5,6} mispairs that are believed to contribute to spontaneous mutations^{1,7–9} and translational errors¹⁰. Chemical modifications that stabilize or lock bases in their anionic or enol-like forms can be mutagenic^{11,12} or expand the decoding capacity of transfer RNAs^{13,14}. In addition, anionic and tautomeric forms of the bases are believed to play crucial roles in nucleic acid catalysis^{15,16}, RNA–ligand recognition^{17,18}, and in the therapeutic mechanisms of nucleic acid base analogues¹⁹.

Despite growing evidence that rare tautomeric and anionic bases have important roles in nucleic acids, their occurrence, stabilities, and biological significance has remained elusive. Characterizing rare tautomeric and anionic bases in polynucleotides is a longstanding problem because such energetically unfavourable species typically exist in low abundance, for short periods of time, and involve movements of protons that are difficult to visualize at the atomic level. NMR relaxation dispersion (RD) techniques^{20–22} enable characterization of low-populated (populations of 0.1–10%) transient (lifetimes of micro- to milliseconds) states of nucleic acids^{23–25} that are often referred to as ‘excited states’ (ES). Here, we use NMR RD to characterize transient WC-like dG•dT and rG•rU mispairs in DNA and RNA that are stabilized by rare tautomeric and anionic bases and obtain evidence that they have universal roles in misincorporation during replication and translation.

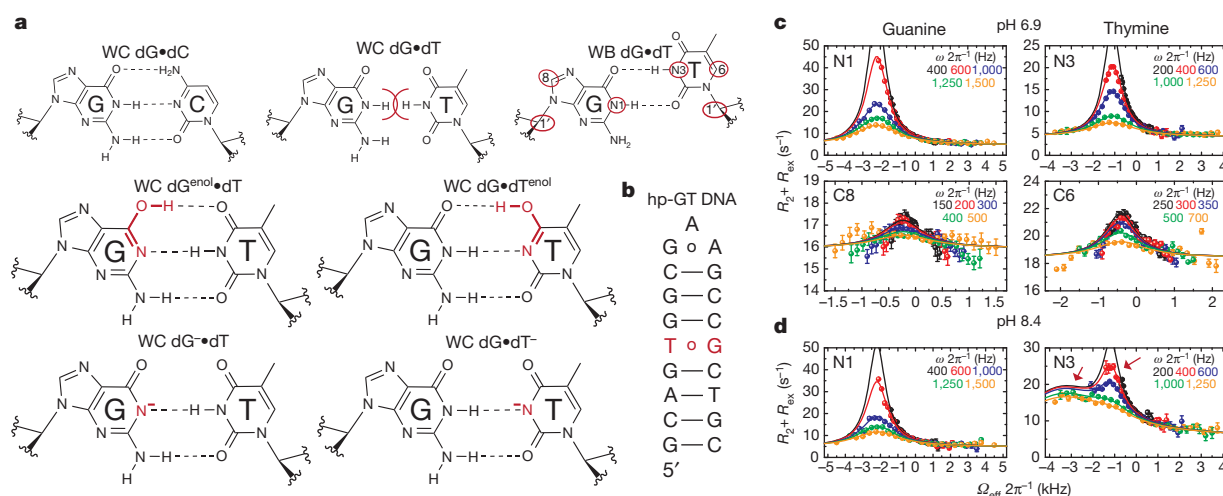


Figure 1 | Chemical exchange in dG•dT mispairs. **a**, WC dG•dC, sterically prohibited WC dG•dT, and WB dG•dT ($R_{1\rho}$ measured nuclei highlighted in red ovals). Below are four WC-like tautomeric and anionic (implied charge delocalization) bp. **b**, DNA duplex with a $^{13}\text{C}/^{15}\text{N}$ site-labelled dG•dT mispair. **c**, RD profiles for dG•dT (25 °C and pH 6.9) showing $R_2 + R_{\text{ex}}$ as a function of the spin lock offset ($\Omega_{\text{eff}} 2\pi^{-1}$) and power ($\omega_{\text{SL}} 2\pi^{-1}$, in insets) with global fits

to dG•N1, dG•C8, dT•N3, and dT•C6. Error bars represent experimental uncertainty (one s.d., see Methods). $\chi^2 = 1.59$, $\bar{R}^2 = 0.99$, and $P < 0.0001$. **d**, RD profiles showing three-state exchange (25 °C and pH 8.4) and global three-state fit to dG•N1 and dT•N3. Error bars represent experimental uncertainty (one s.d.). $\chi^2 = 0.88$, $\bar{R}^2 = 0.99$, and $P < 0.0001$.

¹Department of Biochemistry and Chemistry, Duke University Medical Center, Durham, North Carolina 27710, USA. ²Department of Medical Biochemistry and Biophysics, Karolinska Institute, SE-171 77 Stockholm, Sweden. ³Biophysics, University of Michigan, Ann Arbor, Michigan 48109, USA.

Transient WC-like dG•dT tautomer mispair

dG•dT mispairs generally adopt a distinct ‘wobble’ (WB) geometry (Fig. 1a) since a WC geometry results in a steric clash between imino protons (Fig. 1a). However, enol tautomers of dG or dT, or their anionic form, can alleviate this steric clash and allow formation of WC-like dG•dT mispairs (Fig. 1a). Soon after the discovery of the DNA double helix, Watson and Crick hypothesized that such WC-like mispairs could provide a basis for spontaneous mutations⁷. We used NMR rotating frame spin relaxation ($R_{1\rho}$) RD^{20,26,27} to examine whether WB dG•dT mispairs can transiently morph into such WC-like dG•dT mispairs in canonical DNA duplexes. For these studies we used a hairpin DNA duplex (hp-GT DNA) containing a site-specifically $^{13}\text{C}/^{15}\text{N}$ -labelled dG•dT WB mispair (Fig. 1b and Extended Data Fig. 1). Exchange between WB and WC dG•dT mispairs entails deprotonation of either dG-N1 or dT-N3 via tautomerization (neutral) or ionization (charged), both of which would induce large changes in N1/3 chemical shifts (CSs) and therefore give rise to substantial ^{15}N RD. In contrast, because a WB-to-WC transition preserves an *anti* base and C2'-endo sugar pucker, it is expected to induce smaller changes in the sugar (dG-C1' and dT-C1') and base (dG-C8 and dT-C6) carbon CSs and therefore induce more limited ^{13}C RD. We observed very substantial ^{15}N RD at base imino dG-N1 and to a lesser extent at dT-N3, much less substantial ^{13}C RD at base dG-C8 and dT-C6, and essentially no ^{13}C RD at sugar dG-C1' and dT-C1' at pH 6.9 and 25 °C (Fig. 1c and Extended Data Figs 2, 3). This unique pattern of RD is consistent with exchange directed towards a transient WC-like mispair (Fig. 1a). It is inconsistent with exchange directed towards other base pair (bp) geometries such as Hoogsteen²⁴ or base opened states (Supplementary Discussion 1 and Extended Data Fig. 4). A second exchange process that was apparent at pH 8.4 (Fig. 1d) will be discussed further below. Similar RD profiles were observed in a different DNA duplex (Extended Data Figs 1, 2), indicating that the observed dG•dT exchange occurs robustly in DNA duplexes.

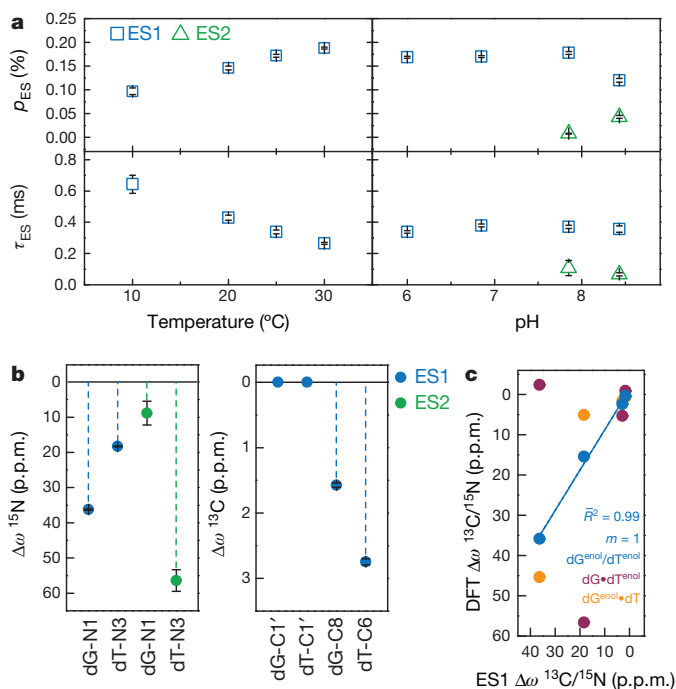


Figure 2 | Characterizing WC-like transient states. **a**, Population and lifetime of dG•dT ES1/ES2 measured in hp-GT DNA as a function of temperature (at pH 6.9) and pH (at 25 °C). Error bars in fitted parameters reflect experimental uncertainty (one s.d.). **b**, Differences between the GS (referenced to 0 p.p.m.) and ES CSs ($\Delta\omega = \omega_{\text{ES}} - \omega_{\text{GS}}$) for hp-GT DNA. Error bars reflect experimental uncertainty (one s.d.). **c**, Measured CSs for ES1 are plotted against DFT-predicted values. $R^2 = 0.99$ and $P = 0.003$.

The RD data measured at dG-N1, dT-N3, dG-C8 and dT-C6 could be globally fitted (Supplementary Table 1 and Extended Data Figs 2, 3) to a single exchange process directed towards an excited state (ES1) that has a population (p_{ES1}) of $\sim 0.17\%$ and a lifetime (τ_{ES1}) of ~ 0.38 ms (Fig. 2a). ES1 is characterized by ^{15}N CSs that are considerably downfield shifted for dG-N1 ($\Delta\omega_{\text{N1}} = +36$ p.p.m.) and to a lesser extent dT-N3 ($\Delta\omega_{\text{N3}} = +18$ p.p.m.) (Fig. 2b, Extended Data Fig. 4). The downfield-shifted imino nitrogen CSs are unprecedented for nucleic acids, and are directed towards the CSs of bases (dG and dT) that have been deprotonated owing to ionization or modifications that lock an enol-like form ($\Delta\omega \approx +50$ – 60 p.p.m.) (Extended Data Fig. 5)^{28–30}. On the other hand, ES1 features much smaller changes in carbon CSs (Fig. 2b), consistent with a WB-to-WC transition.

It would be highly energetically disfavoured to simultaneously deprotonate dG-N1 and dT-N3 when forming ES1. Moreover, although the magnitude of the ES1 N1/N3 downfield CSs strongly suggests deprotonation, it is not as far downfield shifted as expected based on deprotonation of nucleotides in free solution (Extended Data Fig. 5). Therefore, the strongly but incompletely downfield-shifted dG-N1 and dT-N3 CSs suggest that ES1 consists of at least two WC-like species in rapid exchange on the NMR timescale in which either dG-N1 (dG^{enol}•dT or dG•dT⁻) or dT-N3 (dG•dT^{enol} or dG•dT⁻) is deprotonated (Fig. 1a). The ES1 population and CSs are largely independent of pH within the pH range of 6.0–7.9 (Fig. 2a and Supplementary Discussion 2). This is inconsistent with exchange directed towards ionic dG⁻•dT and dG•dT⁻ (Extended Data Fig. 5). Rather, the population of ES1 increases with temperature (Fig. 2a) as expected for tautomeric species dG^{enol}•dT and dG•dT^{enol}.

Based on the measured p_{ES1} and τ_{ES1} , the free energy difference (ΔG) between GS and ES1 is 3.8 kcal mol⁻¹ and the forward free energy barrier (ΔG^\ddagger) is 16.4 kcal mol⁻¹ (Extended Data Fig. 6). These values are in good agreement with computationally predicted parameters (2.8 – 5.6 kcal mol⁻¹ (ref. 31) and ~ 17 – 21 kcal mol⁻¹ (refs 31, 32), respectively) for dG•dT WB-to-WC tautomer transitions.

These computational studies also predict that dG^{enol}•dT probably exists in fast exchange on the NMR timescale (free energy barrier ~ 5 – 6 kcal mol⁻¹, ref. 31) with a minor (20%) dG•dT^{enol} species (Fig. 1a and Supplementary Discussions 3, 4). Under these conditions, the measured ES1 CSs would represent a population-weighted average of the two tautomeric states (Methods). We find that all ES1 CSs (dG-N1, dT-N3, dG-C8 and dT-C6) are in quantitative agreement with values predicted by density functional theory (DFT)³³ calculations for a weighted dG^{enol}•dT(80%) \rightleftharpoons dG•dT^{enol}(20%) equilibrium (Figs 2c and 3a).

Transient WC-like dG•dT⁻ anionic mispair

Interestingly, upon increasing the pH to 8.4, we observed evidence for a second excited state (ES2), which is seen as a second peak in the off-resonance RD profile of dT-N3 (Fig. 1d). Global fitting of these RD data (Extended Data Figs 2 and 3) revealed two excited states (ES1 and ES2) that are most probably arranged in a linear topology (ES1 \rightleftharpoons dG•dT \rightleftharpoons ES2).

Compared to ES1, ES2 ($\Delta G^\ddagger = 16.4$ and $\Delta G = 4.59$ kcal mol⁻¹) has a considerably lower population ($p_{\text{ES2}} \approx 0.04\%$) and lifetime ($\tau_{\text{ES2}} \approx 70$ μs) at pH 8.4 (Fig. 2a). The dG-N1 and dT-N3 ES2 CSs are not only ‘swapped’ relative to ES1 such that dT-N3 experiences the larger downfield shift ($\Delta\omega_{\text{N3}} = +56$ p.p.m.) while dG-N1 experiences a smaller downfield shift ($\Delta\omega_{\text{N1}} = +9$ p.p.m.) (Fig. 2b), they are also more asymmetric in favour of a deprotonated dT species. In addition, unlike p_{ES1} , p_{ES2} increases considerably with pH, consistent with ionization and the formation of dG•dT⁻ (Fig. 3a). The ES2 CSs are in excellent agreement with values computed using DFT assuming a WC-like dG•dT⁻ (100%) species (Fig. 3a, b). However, we cannot rule out that dG•dT⁻ is in rapid equilibrium with a WC-like dG⁻•dT or an inverted wobble (iWB) geometry^{3,34,35} (Extended Data Fig. 4 and Supplementary Discussion 1) that falls outside detection limits (Fig. 3a).

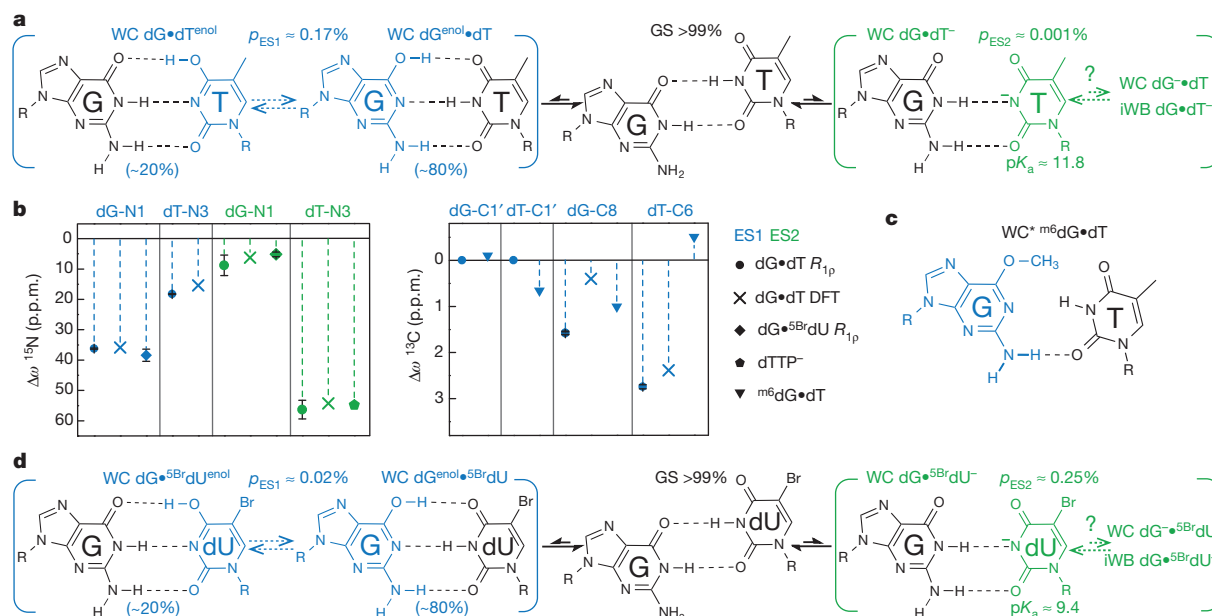


Figure 3 | Mutate-and-CS fingerprinting ES1 and ES2. **a**, Multi-state equilibrium between WB and WC-like dG•dT mispairs. ES1/ES2 populations and weights are shown (25 °C and pH 6.9). p_{ES2} estimated based on the observed apparent pK_a . **b**, CS fingerprinting dG•dT ES1/ES2 using chemical modifications and structure-based DFT predictions of CSs. ES1 DFT CSs

are given for 80:20 dG^{enol}•dT:dG•dT^{enol} weighting. Error bars reflect experimental uncertainty (one s.d.). **c**, ^{m6}dG•dT structure^{12,30}. **d**, dG•^{5Br}dU ES2 stabilized relative to dG•dT^{enol} ES2. Populations and weights (10 °C and pH 6.9) are shown.

Fingerprinting the dG•dT excited states

We adapted a mutate-and-CS fingerprint strategy^{24,25} to test the proposed ES1 and ES2 (Fig. 3a); chemical modifications are used to trap an ES, or induce specific perturbations to the GS \rightleftharpoons ES equilibrium. We trapped ES1 (dG^{enol}•dT) using the mutagenic base O6-methyl-2'-deoxyguanosine (^{m6}dG) (Fig. 3c and Extended Data Fig. 7), which is known to adopt a distorted WC dG^{enol}•dT-like mispair^{12,30}. Relative to the WB, this modification resulted in negligible changes in dG-C1' ($\Delta\omega_{C1'} = -0.1$ p.p.m.) and dT-C1' ($\Delta\omega_{C1'} = +0.7$ p.p.m.) CSs, and a small downfield shift in dG-C8 ($\Delta\omega_{C8} = +1$ p.p.m.), consistent with the RD-derived ES1 CSs (Fig. 3b and Extended Data Fig. 7). The modification induced a small upfield shift in dT-C6 ($\Delta\omega_{C6} = -0.5$ p.p.m.) that is inconsistent with the downfield dT-C6 CS (Fig. 3b) observed by RD. However, such a deviation is expected on the basis of DFT calculations ($\Delta\omega_{C6} = -2.3$ p.p.m. for dT-C6 in the ^{m6}dG•T pair) (Supplementary Discussion 5) and can be attributed to minor deviations from an ideal dG^{enol}•dT WC-like mispair geometry (Fig. 3c)^{12,30}. Severe line broadening did not permit measurement of the dT-N3 CSs in these non-isotopically enriched samples.

To test the proposed dG•dT^{enol} ES2, we measured the difference in dT-N3 CS between neutral and anionic dTTP ($\Delta\omega_{N3} = +55$ p.p.m.) and found them to be in excellent agreement with the dT-N3 CS differences measured by RD ($\Delta\omega_{N3} = +56$ p.p.m.) (Fig. 3b and Extended Data Fig. 5). In addition, we used the mutagenic thymidine-analogue, 5-bromo-2'-deoxyuridine (^{5Br}dU) to push the equilibrium towards dG•^{5Br}dU^{enol} (Fig. 3d). This modification lowers the pK_a of ^{5Br}dU-N3 (~8.6) and favours a WC-like dG•^{5Br}dU^{enol} geometry at high pH¹¹. This modification increased the population of ES2 ($\Delta G^\ddagger = 15.1$ and $\Delta G = 3.37$ kcal mol⁻¹) by over two orders of magnitude at the expense of ES1 ($\Delta G^\ddagger = 16.0$ and $\Delta G = 4.67$ kcal mol⁻¹) while minimally affecting the ES1 and ES2 CSs (Fig. 3b). The consistencies in ES2 CSs between dG•dT^{enol} and dG•^{5Br}dU^{enol} mispairs further support a dominant WC-like ES2, rather than an iWB bp, in naked DNA. The unaffected ES1 CSs indicate that ^{5Br}dU does not appreciably impact the dG^{enol}•^{5Br}dU \rightleftharpoons dG•^{5Br}dU^{enol} equilibrium (Fig. 3d) relative to dG•dT (Fig. 3a), consistent with prior computational studies^{31,36}.

Transient WC-like rG•rU mispairs in RNA

If the observed ESs correspond to WC-like dG•dT mispairs, it is predicted that similar ESs should arise in rG•rU mispairs in RNA where WC base pairs are also readily accommodated within the A-form helix. To test this hypothesis, we carried out analogous pH- and temperature-dependent RD measurements on two RNA duplexes. RD profiles measured for rG•rU in A-form RNA (Fig. 4a and Extended Data Fig. 8) are very similar to those measured for dG•dT in B-form DNA (Fig. 1c, d). Global analyses of the RD data (Supplementary Table 1) revealed an apparent three-state exchange process at pH ≥ 7.9 (Extended Data Fig. 3). The RD-derived CSs (Fig. 4b and Extended Data Fig. 9), together with the pH and temperature dependence of the populations and lifetimes (Fig. 4c), are consistent with rG•rU^{enol} \rightleftharpoons rG^{enol}•rU as ES1 and rG•rU^{enol} as ES2 (Fig. 4d) as observed in DNA. The rG•rU ES1 forward free energy barrier ($\Delta G^\ddagger = 15.7$ and $\Delta G = 3.86$ kcal mol⁻¹) is consistent with barriers measured for dG•dT ES1 (16.4 kcal mol⁻¹) and computationally predicted barriers for G•U tautomerization (17.1 kcal mol⁻¹)³². Computational studies³⁶ show that dG•dU^{enol} is stabilized relative to dG•dT^{enol}. This is predicted to tilt the rapid rG^{enol}•rU \rightleftharpoons rG•rU^{enol} equilibrium in favour of rG•rU^{enol} (40%) in RNA as compared to dG•dT^{enol} (20%) in DNA (Methods). We find that in RNA, the ES1 rU-N3 CSs are slightly more downfield shifted ($\Delta\omega_{N3} = +30$ p.p.m.) than rG-N1 ($\Delta\omega_{N1} = +26$ p.p.m.) (Fig. 4b). Reweighting the DFT-predicted CSs assuming 60:40 ratios of rG^{enol}•rU:rG•rU^{enol} gives an excellent fit to RD-derived values (Fig. 4b), and are in better agreement than 80:20 rG^{enol}•rU:rG•rU^{enol} (Extended Data Fig. 9). Changes observed in the ES1 CSs at higher pH suggest a potentially more complex exchange process (Supplementary Discussion 6). As with dG•dT^{enol}, the WC-like rG•rU^{enol} may exist in equilibrium with both a WC-like rG^{enol}•rU and/or an iWB rG•rU^{enol} (Fig. 4d).

rG•rU wobbles are widespread in RNA, where they play important structural and functional roles³⁷. We therefore examined whether the ESs observed here would occur in more complex RNA structural contexts. Indeed, the rU-N3 ¹⁵N RD profiles measured for two WB rG•rU mispairs in a 69-nucleotide *Bacillus subtilis* guanine riboswitch (Extended Data Fig. 8) provide initial evidence ($\Delta\omega_{N3} = +44$ –47 p.p.m.

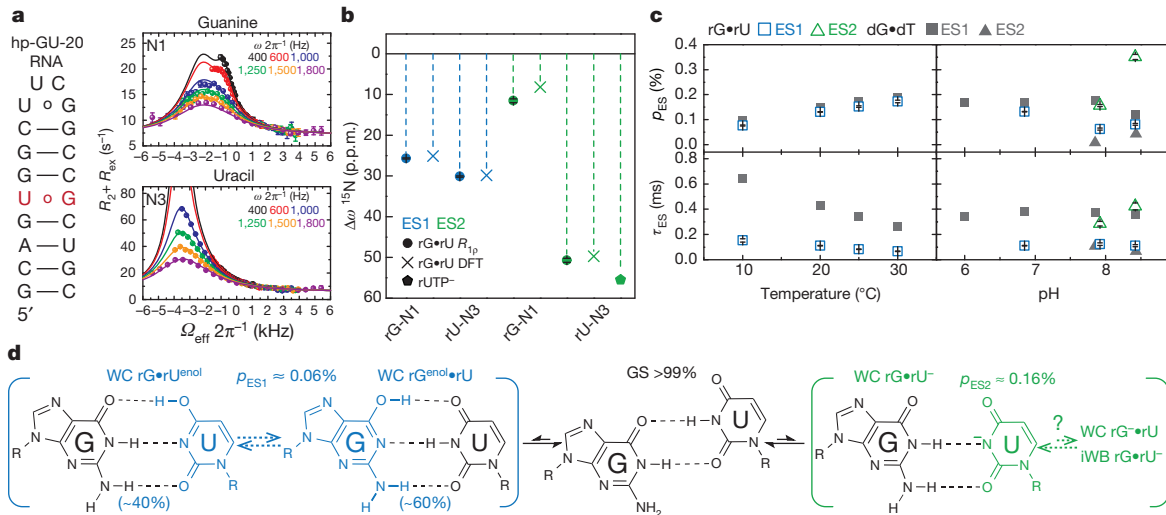


Figure 4 | Transient tautomeric and anionic WC-like mispairs in A-form RNA. **a**, RNA duplex and $R_{1\rho}$ RD profiles with three-state global fits to rG-N1 and rU-N3 (20 °C and pH 7.9). Error bars represent experimental uncertainty (one s.d.). $\chi^2 = 0.66$, $R^2 = 0.99$, and $P < 0.0001$. **b**, CSs for rG•rU ES1/ES2 compared to structure-based DFT predictions and rUTP ionization. Error bars

reflect experimental uncertainty (one s.d.). **c**, Population and lifetime of rG•rU ES1/ES2 measured as a function of temperature (at pH 6.9) and pH (at 20 °C). dG•dT ES1 and ES2 shown (in grey) for comparison. Error bars reflect experimental uncertainty (one s.d.). **d**, Multi-state equilibrium between WB and WC-like rG•rU mispairs (20 °C and pH 7.9).

and $p_{ES} \approx 0.04\%$ at pH 7.9) for WC-like rG•rU[−] mispairs in more complex RNA structures (Extended Data Fig. 9 and Supplementary Discussion 7). Therefore, we expect that transient WC-like rG•rU mispairs exist robustly across the RNA transcriptome.

Biological implications

Initial selection of NTPs during replication, and tRNAs during translation, strongly relies on WC stereochemical geometry as a means of discriminating against mispairs^{38,39}. The low error rate ($\sim 10^{-3}$ to 10^{-6}) during initial selection accounts for most of the overall fidelity of replication ($\sim 10^{-6}$ to 10^{-10})^{40–42} and translation ($\sim 10^{-3}$ to 10^{-5})^{43–45}. By stereochemically mimicking the WC geometry, the ES WC-like dG•dT and rG•rU mispairs observed here can provide a mechanism for evading WC fidelity checks during initial substrate selection^{1,8,10}. The intrinsic probabilities with which WC-like mispairs form have long been suspected as important determinants of substitution mutation probability^{8,10}. By carrying out the first measurements of the intrinsic probabilities with which WC-like dG•dT and rG•rU mispairs form in native DNA/RNA systems, we are able to obtain unique insights into the mechanisms of misincorporation and the potential roles of ES1/ES2.

We find that the probabilities with which WC-like dG•dT ES1 and ES2 form in duplex DNA (10^{-3} to 10^{-5}) span the dGTP•dT/dG•dTTP misincorporation and base substitution probabilities seen during replication using high-fidelity polymerases with little/no proofreading capabilities (Fig. 5a, Extended Data Fig. 10 and Supplementary Discussion 8)^{46–48}. Similarly, the WC-like rG•rU ES1 and ES2 probabilities (10^{-3} to 10^{-4}) span the majority of amino acid misincorporation probabilities arising specifically due to rG•rU pairing at any codon position (10^{-3} to 10^{-5})^{43,44} (Fig. 5b and Supplementary Discussion 8). Note that some of the amino acid misincorporation probabilities (10^{-5}) are lower than the measured rG•rU ES2 probability, which could be due to translational proofreading⁴⁵ and/or lower pH conditions that destabilize ES2. These results, together with previous structural studies showing that WB and WC-like mispairs can exist within polymerase^{1–3} and ribosome^{5,6,13} active sites, strongly suggest that energetic competition between WB and WC-like mispairs is robust and is a key determinant of misincorporation probability during replication and translation (Supplementary Discussion 9). However, additional studies are needed to determine the probabilities with which WC-like mispairs form within the unique environment of polymerases and ribosomes. A recent MD study found that rG^{enol}•rU tautomers can be stabilized in a ribosome

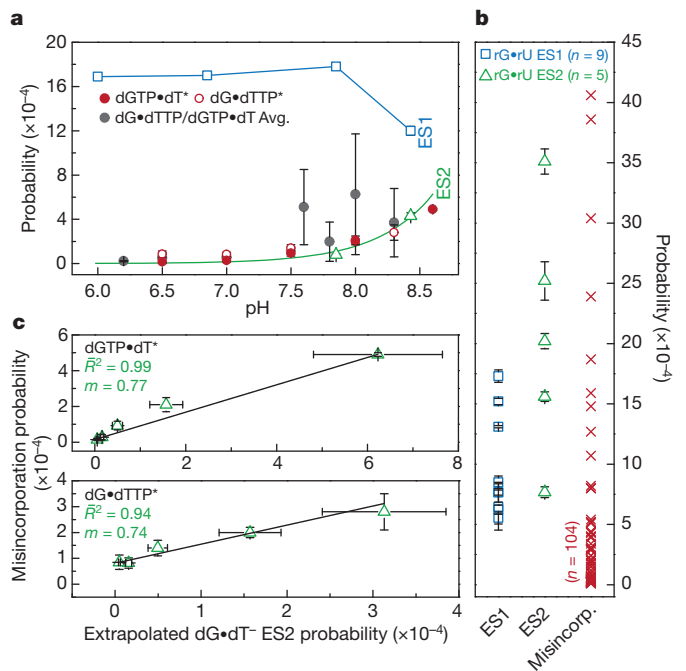


Figure 5 | Correlation between WC-like bp and misincorporation probabilities. **a**, dG•dT ES1 (blue square) and ES2 (green triangle) probabilities; error bars reflect experimental uncertainty (one s.d.). pH-dependent dGTP•dT*/dG•dTTP* misincorporation probabilities (errors as published)⁴⁷. Averaged dGTP•dT/dG•dTTP misincorporation and base substitution probabilities; error bars reflect the s.d. (Supplementary Discussion 8). ES2 fit to Henderson–Hasselbalch equation shown as green line (see Methods). A blue line connects ES1 points for visualization purposes. **b**, rG•rU ES1 (blue square, $n = 9$) and ES2 (green triangle, $n = 5$) probabilities measured at varied conditions (Supplementary Table 1), and amino acid misincorporation probabilities (red 'X', $n = 104$) due to rG•rU mispairs as given in ref. 44. Error bars in ES1/ES2 reflect experimental uncertainty (one s.d.). **c**, The probabilities of dGTP•dT/dG•dTTP misincorporation (error as given) by avian myeloblastosis virus RT (ref. 47) versus the pK_a -predicted probability of forming a WC-like dG•dT[−] mispair as a function of pH; error bars reflect experimental uncertainty (one s.d.). dGTP•dT $R^2 = 0.99$ and $P = 0.0001$. dG•dTTP $R^2 = 0.94$ and $P = 0.004$.

context, but challenges their involvement in decoding errors based on predicted tRNA binding energies⁴⁹.

The pH-dependent dG•dT misincorporation probability^{47,48,50} points to the involvement of an anionic species in misincorporation^{1,47}. Our results strongly suggest that this species is most likely to be ES2 dG•dT[−] and not the energetically disfavoured dG••dT. We observe excellent agreement between the pK_a-predicted probability of forming ES2 dG•dT[−] and pH-dependent dGTP•dT/dG•dTTP misincorporation probabilities measured for a reverse transcriptase⁴⁷ that lacks any proofreading ability (Fig. 5c and Extended Data Fig. 10). We note that the correlation is reduced at more extreme pH, near the pK_as of other ionizable groups in proteins and DNAs (Extended Data Fig. 10 and Supplementary Discussion 10). We also find that dG•^{5Br}dU[−] enhances the probability of forming a WC-like ES2 (Fig. 3d) and correspondingly results in an approximately eightfold increase in dG•^{5Br}dU misincorporation⁴⁷. These data indicate that, for this polymerase, misincorporation proceeds predominantly via a WC-like dG•dT[−] as proposed⁴⁷. It is very likely that parameters such as polymerase types, DNA sequence, and the chemical environment can affect the relative stabilities and lifetimes of the anionic, tautomeric, and WB mispairs³. Therefore we expect this to affect the flux through distinct misincorporation pathways involving different WC and WB conformations, which may help to explain the broad range of misincorporation probabilities.

Our findings suggest that unconventional enol tautomeric and anionic bases exist robustly in genomes. We anticipate that these rare tautomeric and anionic bases play unique roles in DNA damage induction and repair, nucleic acid recognition, chemical modifications of nucleic acids, and catalysis. The NMR methods outlined here can immediately be applied to characterize tautomeric and anionic species, which we believe will not be restricted to dG•dT and rG•rU mispairs, but rather will be widespread across diverse nucleic acid motifs.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 23 October 2014; accepted 9 January 2015.

Published online 11 March 2015.

- Bebenek, K., Pedersen, L. C. & Kunkel, T. A. Replication infidelity via a mismatch with Watson–Crick geometry. *Proc. Natl Acad. Sci. USA* **108**, 1862–1867 (2011).
- Koag, M.-C., Nam, K. & Lee, S. The spontaneous replication error and the mismatch discrimination mechanisms of human DNA polymerase β . *Nucleic Acids Res.* **42**, 11233–11245 (2014).
- Xia, S. & Konigsberg, W. H. Mispairs with Watson–Crick base-pair geometry observed in ternary complexes of an RB69 DNA polymerase variant. *Protein Sci.* **23**, 508–513 (2014).
- Wang, W., Hellings, H. W. & Beese, L. S. Structural evidence for the rare tautomer hypothesis of spontaneous mutagenesis. *Proc. Natl Acad. Sci. USA* **108**, 17644–17648 (2011).
- Ogle, J. M., Murphy, F. V., Tarry, M. J. & Ramakrishnan, V. Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* **111**, 721–732 (2002).
- Demeshkina, N., Jenner, L., Westhof, E., Yusupov, M. & Yusupova, G. A new understanding of the decoding principle on the ribosome. *Nature* **484**, 256–259 (2012).
- Watson, J. D. & Crick, F. H. C. The structure of DNA. *Cold Spring Harb. Symp. Quant. Biol.* **18**, 123–131 (1953).
- Topal, M. D. & Fresco, J. R. Complementary base pairing and the origin of substitution mutations. *Nature* **263**, 285–289 (1976).
- Harris, V. H. et al. The effect of tautomeric constant on the specificity of nucleotide incorporation during DNA replication: support for the rare tautomer hypothesis of substitution mutagenesis. *J. Mol. Biol.* **326**, 1389–1401 (2003).
- Topal, M. D. & Fresco, J. R. Base pairing and fidelity in codon–anticodon interaction. *Nature* **263**, 289–293 (1976).
- Sowers, L. C., Goodman, M., Eritja, R., Kaplan, B. & Fazakerley, G. Ionized and wobble base-pairing for bromouracil–guanine in equilibrium under physiological conditions: A nuclear magnetic resonance study on an oligonucleotide containing a bromouracil–guanine base-pair as a function of pH. *J. Mol. Biol.* **205**, 437–447 (1989).
- Warren, J. J., Forsberg, L. J. & Beese, L. S. The structural basis for the mutagenicity of O⁶-methyl–guanine lesions. *Proc. Natl Acad. Sci. USA* **103**, 19701–19706 (2006).
- Weixlbaumer, A. et al. Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. *Nature Struct. Mol. Biol.* **14**, 498–502 (2007).
- Cantara, W. A., Murphy, F. V., Demirci, H. & Agris, P. F. Expanded use of sense codons is regulated by modified cytidines in tRNA. *Proc. Natl Acad. Sci. USA* **110**, 10964–10969 (2013).
- Bevilacqua, P. C. & Yajima, R. Nucleobase catalysis in ribozyme mechanism. *Curr. Opin. Chem. Biol.* **10**, 455–464 (2006).
- Cochrane, J. C. & Strobel, S. A. Catalytic strategies of self-cleaving ribozymes. *Acc. Chem. Res.* **41**, 1027–1035 (2008).
- Gilbert, S. D., Reyes, F. E., Edwards, A. L. & Batey, R. T. Adaptive ligand binding by the purine riboswitch in the recognition of guanine and adenine analogs. *Structure* **17**, 857–868 (2009).
- Singh, V. et al. Direct observation of multiple tautomers of oxythiamine and their recognition by the thiamine pyrophosphate riboswitch. *ACS Chem. Biol.* **9**, 227–236 (2014).
- Li, D. et al. Tautomerism provides a molecular explanation for the mutagenic properties of the anti-HIV nucleoside 5-aza-5,6-dihydro-2'-deoxycytidine. *Proc. Natl Acad. Sci. USA* **111**, E3252–E3259 (2014).
- Korzhnev, D. M., Orekhov, V. Y. & Kay, L. E. Off-resonance R_{1ρ} NMR studies of exchange dynamics in proteins with low spin-lock fields: an application to a Fyn SH3 domain. *J. Am. Chem. Soc.* **127**, 713–721 (2005).
- Palmer, A. G. III. Chemical exchange in biomacromolecules: past, present, and future. *J. Magn. Reson.* **241**, 3–17 (2014).
- Hansen, A. L. & Kay, L. E. Measurement of histidine pK_a values and tautomer populations in invisible protein states. *Proc. Natl Acad. Sci. USA* **111**, E1705–E1712 (2014).
- Hoogstraten, C. G., Wank, J. R. & Pardi, A. Active site dynamics in the lead-dependent ribozyme. *Biochemistry* **39**, 9951–9958 (2000).
- Nikolova, E. N. et al. Transient Hoogsteen base pairs in canonical duplex DNA. *Nature* **470**, 498–502 (2011).
- Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negroni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature* **491**, 724–728 (2012).
- Massi, F., Johnson, E., Wang, C., Rance, M. & Palmer, A. G. III. NMR R_{1ρ} rotating-frame relaxation with weak radio frequency fields. *J. Am. Chem. Soc.* **126**, 2247–2256 (2004).
- Hansen, A. L., Nikolova, E. N., Casiano-Negroni, A. & Al-Hashimi, H. M. Extending the range of microsecond-to-millisecond chemical exchange detected in labeled and unlabeled nucleic acids by selective carbon R_{1ρ} NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 3818–3819 (2009).
- Duthaler, R. O. & Roberts, J. D. Effects of solvent, protonation, and N-alkylation on the nitrogen-15 chemical shifts of pyridine and related compounds. *J. Am. Chem. Soc.* **100**, 4969–4973 (1978).
- Cho, B. P., Kadlubar, F. F., Culp, S. J. & Evans, F. E. Nitrogen-15 nuclear magnetic resonance studies on the tautomerism of 8-hydroxy-2'-deoxyguanosine, 8-hydroxyguanosine, and other C8-substituted guanine nucleosides. *Chem. Res. Toxicol.* **3**, 445–452 (1990).
- Goswami, B., Gaffney, B. L. & Jones, R. A. Nitrogen-15-labeled oligodeoxynucleotides. 5. Use of ¹⁵N NMR to probe H-bonding in an O⁶MeG•T base pair. *J. Am. Chem. Soc.* **115**, 3832–3833 (1993).
- Nomura, K. et al. DFT calculations on the effect of solvation on the tautomeric reactions for wobble Gua–Thy and canonical Gua–Cyt base-pairs. *J. Mod. Phys.* **4**, 422–431 (2013).
- Brovarets', O. O. Effect of a modification of uracil on the tautomerization barrier of the wobble Gua•^{5X}Ura base pair into the Gua*•^{5X}Ura base pair with the Watson–Crick geometry: quantum-chemical study. *Rep. Natl. Acad. Sci. Ukraine* **4**, 154–158 (2013).
- Xu, X.-P. & Au–Yeung, S. C. F. Investigation of chemical shift and structure relationships in nucleic acids using NMR and density functional theory methods. *J. Phys. Chem. B* **104**, 5641–5650 (2000).
- Johnson, S. J. & Beese, L. S. Structures of mismatch replication errors observed in a DNA polymerase. *Cell* **116**, 803–816 (2004).
- Brovarets', O. O., Zhuravskiy, R. O. & Hovorun, D. M. Is there adequate ionization mechanism of the spontaneous transitions? Quantum-chemical investigation. *Biopolymers Cell* **26**, 398–405 (2010).
- Orozco, M., Hernández, B. & Luque, F. J. Tautomerism of 1-methyl derivatives of uracil, thymine, and 5-bromouracil. Is tautomerism the basis for the mutagenicity of 5-bromouridine? *J. Phys. Chem. B* **102**, 5228–5233 (1998).
- Varani, G. & McClain, W. H. The G•U wobble base pair. *EMBO Rep.* **1**, 18–23 (2000).
- Kunkel, T. A. & Bebenek, K. DNA replication fidelity. *Annu. Rev. Biochem.* **69**, 497–529 (2000).
- Voorhees, R. M. & Ramakrishnan, V. Structural basis of the translational elongation cycle. *Annu. Rev. Biochem.* **82**, 203–236 (2013).
- Kunkel, T. A. DNA replication fidelity. *J. Biol. Chem.* **279**, 16895–16898 (2004).
- Schaaper, R. M. Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli*. *J. Biol. Chem.* **268**, 23762–23765 (1993).
- Xia, S. & Konigsberg, W. H. RB69 DNA polymerase structure, kinetics, and fidelity. *Biochemistry* **53**, 2752–2767 (2014).
- Ogle, J. M. & Ramakrishnan, V. Structural insights into translational fidelity. *Annu. Rev. Biochem.* **74**, 129–177 (2005).
- Zhang, Z., Shah, B. & Bondarenko, P. V. G/U and certain wobble position mismatches as possible main causes of amino acid misincorporations. *Biochemistry* **52**, 8165–8176 (2013).
- Gromadski, K. B. & Rodnina, M. V. Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol. Cell* **13**, 191–200 (2004).
- Kunkel, T. A. & Alexander, P. S. The base substitution fidelity of eucaryotic DNA polymerases. Mismatching frequencies, site preferences, insertion preferences, and base substitution by dislocation. *J. Biol. Chem.* **261**, 160–166 (1986).

47. Yu, H., Eritja, R., Bloom, L. B. & Goodman, M. F. Ionization of bromouracil and fluorouracil stimulates base mispairing frequencies with guanine. *J. Biol. Chem.* **268**, 15935–15943 (1993).
48. Eckert, K. A. & Kunkel, T. A. Effect of reaction pH on the fidelity and processivity of exonuclease-deficient Klenow polymerase. *J. Biol. Chem.* **268**, 13462–13471 (1993).
49. Satpati, P. & Åqvist, J. Why base tautomerization does not cause errors in mRNA decoding on the ribosome. *Nucleic Acids Res.* **42**, 12876–12884 (2014).
50. Driggers, P. H. & Beattie, K. L. Effect of pH on the base-mispairing properties of 5-bromouracil during DNA synthesis. *Biochemistry* **27**, 1729–1735 (1988).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank S. Horowitz, H. Zhou, J. Lee, A. M. Mustoe, and E. N. Nikolova for assistance and critical comments. We are grateful for technical support and resources from the Duke Magnetic Resonance Spectroscopy Center and

University of Michigan Flux HPC Cluster. This work was supported by an NIH grant (R01GM089846) and an Agilent Thought Leader Award given to H.M.A.

Author Contributions I.J.K. and H.M.A. conceived the project and experimental design. I.J.K. prepared NMR samples as well as performed and analysed all NMR RD experiments. I.J.K. assigned resonances in all nucleic acid constructs with assistance from B.S.; K.P. prepared the hp-GU-24 sample and carried out additional NMR RD experiments. I.J.K. performed all DFT calculations. Z.W.S. assisted I.J.K. with numerical Bloch–McConnell simulations. I.J.K. and H.M.A. wrote the manuscript with critical input from B.S. and K.P.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to H.M.A. (hashim.al.hashimi@duke.edu).

METHODS

Sample preparation

NMR buffer. All duplex DNA and RNA samples were buffer-exchanged using a centrifugal concentrator (EMD Millipore) into a solution containing 25 mM sodium chloride (100 mM sodium chloride for Dickerson-GT DNA), 15 mM sodium phosphate, 0.1 mM EDTA, and 10% D₂O with variable pH (6.0, 6.4, 6.8, 6.9, 7.9). pH values of 8.4 were obtained for hp-GT DNA and hp-GU-20 RNA samples by direct titration of pH 7.9 samples with filtered 0.5 M NaOH solution. Monovalent ion concentration subsequently increased by a small amount proportional to the NaOH titrated in but did not affect DNA/RNA conformation as confirmed by NMR. Natural isotopic abundance oligonucleotide sample concentrations ranged from 2 to 3.5 mM. ¹³C/¹⁵N-labelled oligonucleotide sample concentrations ranged from 0.7 to 3.5 mM. xpt-G riboswitch sample was diluted to ~30 μM in a solution containing saturated guanine, denatured, and annealed on ice. Sample was buffer exchanged against either potassium acetate (pH 6.7) or sodium phosphate (pH 7.9) buffer and concentrated to 0.7–1.7 mM. Mg²⁺ was titrated in until total concentration was ~5 mM.

Site-specifically ¹³C/¹⁵N-labelled DNA samples. Selectively ¹³C/¹⁵N-labelled DNA samples (hp-GT DNA, ⁵BrU5-hp-GT and ⁸BrG15-hp-GT DNA) were purchased from the Yale Keck Oligonucleotide Synthesis Facility and were synthesized using commercially available 2'-deoxyguanosine DMT-phosphoramidite (98% ¹³C10, 98% ¹⁵N5) and 2'-deoxythymidine phosphoramidite (98% ¹³C10, 98% ¹⁵N2) purchased from Cambridge Isotope Labs. hp-GT DNA was selectively ¹³C/¹⁵N-labelled at dT5 and dG15. ⁵BrU5-hp-GT DNA was ¹³C/¹⁵N-labelled at dG15, and ⁸BrG15-hp-GT was ¹³C/¹⁵N-labelled at dT5. Samples were purified using RP-HPLC before buffer exchange. NMR experiments were used to confirm native folding of hp-GT, ⁵BrU5-hp-GT, and ⁸BrG15-hp-GT DNA constructs.

Enzymatic synthesis of ¹³C/¹⁵N-labelled DNA samples. The Dickerson-GT DNA sample was synthesized *in vitro* using uniformly enriched ¹³C/¹⁵N dGTP and dTTP (Silantes GmbH) as previously described⁵¹. Reaction mixture was centrifuged and filtered to remove excess pyrophosphate and concentrated down to 1 ml in a 3 kDa molecular weight cut-off centrifugal concentrator (EMD Millipore). Sample was mixed with 1 ml of a formamide-based denaturing loading dye, denatured at 95 °C for 5 min, and loaded onto a 33 × 102 cm sequencing gel (20% polyacrylamide/8M urea) and run for 12 h to resolve target oligonucleotide from template and other nucleic acid species. Target band was shadowed briefly using a UV hand-lamp and excised before gel electroelution (Whatman, GE Healthcare), followed by ethanol precipitation. Sample purity was confirmed using gel electrophoresis (20% polyacrylamide/8M urea) stained with SybrGOLD before buffer exchange.

Enzymatic synthesis of ¹³C/¹⁵N-labelled RNA samples. The hp-GU-20, hp-GU-24 and xpt-G riboswitch RNA samples were prepared using *in vitro* transcription as previously described²⁵ using uniformly enriched ¹³C/¹⁵N ribonucleotide triphosphates (hp-GU-20 RNA: rGTP & rUTP only, hp-GU-24: All and xpt-G: All). Purification was carried out as described above for ¹³C/¹⁵N-labelled DNA construct.

Unlabelled and unlabelled-modified DNA samples. hp-GT and Dickerson-GT DNA constructs at natural isotopic abundance were purchased from Integrated DNA Technologies. The O6-methyl-2'-deoxyguanosine mismatch constructs (^{m6}G15-hp-GT and ^{m6}G4-Dickerson-GT DNA) were purchased from the Yale Keck Oligonucleotide Synthesis Facility. hp-GT and Dickerson-GT DNA constructs were desalted before buffer exchange. Unlabelled-modified samples were purified using RP-HPLC before buffer exchange.

Isotopically enriched dNTP and rNTP samples. Uniformly ¹³C/¹⁵N enriched dGTP, dTTP, rGTP and rUTP samples were purchased (Silantes GmbH) and added to an NMR buffer (25 mM sodium chloride, 15 mM sodium phosphate, 0.1 mM EDTA and 10% D₂O at pH 6.9). Sample was adjusted to pH ≈ 12.5 directly using 5 M NaOH.

NMR experiments. Resonance assignment. The CS assignment for DNA and RNA constructs were obtained using aromatic [¹³C, ¹H], aliphatic [¹³C, ¹H], imino [¹⁵N, ¹H] heteronuclear and [¹H, ¹H] nuclear Overhauser effect spectroscopy (NOESY) homonuclear correlation experiments. The data for all DNA constructs were acquired on an 18.8 T Agilent spectrometer equipped with a triple resonance HCN cryogenic probe, for the uniformly ¹³C/¹⁵N-labelled hp-GU-24 RNA construct on a 14.1 T Bruker Avance spectrometer equipped with a triple-resonance HCN cryogenic probe, and for the xpt-G RNA riboswitch on a 14.1 T Agilent spectrometer equipped with a Bruker HCPN cryogenic probe. All data were processed and analysed using the software NMRpipe⁵² and SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco), respectively. Resonance assignment for exchangeable and non-exchangeable ¹H was performed using the 2D [¹H, ¹H] WATERGATE NOESY experiment⁵³ (mixing time 250 ms) as described previously^{27,54}, and their covalently bonded ¹³C/¹⁵N nuclei were assigned using heteronuclear single/multiple quantum coherence correlation experiments (HSQC or HMQC). For the labelled RNA constructs conventional HSQC experiments⁵³ were acquired for all spins, while for unlabelled DNA constructs conventional HSQC

was used for the aliphatic C1' spins and band-selective optimized flip angle short-transient (SOFAT) HMQC^{55,56} were employed for the imino and aromatic spins.

¹⁵N R_{1ρ} relaxation dispersion. 1D ¹⁵N R_{1ρ} RD experiments^{27,57} targeting imino nitrogen resonances of interest were carried out at 14.1 T (hp-GT, ⁵BrU5-hp-GT, Dickerson-GT DNA and hp-GU-24 RNA) and 16.4 T (hp-GU-20 RNA) as previously described⁵⁷. Raw data were processed using NMRpipe⁵² to generate a series of peak intensities. On- and off-resonance R_{1ρ} RD profiles were recorded using spinlock powers (ω_{SL} 2π⁻¹) ranging from 100 to 2,000 Hz, with the absolute offset frequencies (Ω 2π⁻¹ Hz) ranging from 0–3.5× the applied spinlock power (Supplementary Table 1). Offset frequencies greater than 3.5× the given spinlock power were not used owing to substantial R₁ relaxation contributions²⁷. Magnetization of the spins of interest was allowed to relax under an applied spinlock for the following durations: 0–120 ms for N1/N3 of hp-GT, ⁵BrU5-hp-GT, Dickerson-GT DNA and hp-GU-24 RNA; 0–100 ms for N1/N3 of hp-GU-20 RNA; and 0–80 ms, 0–74 ms, 0–68 ms for N3 of the xpt-G riboswitch.

¹³C R_{1ρ} relaxation dispersion. ¹³C R_{1ρ} RD experiments targeting carbon resonances of interest were carried out at 14.1 T as previously described^{24,27}. On- and off-resonance R_{1ρ} RD profiles were measured using spinlock powers (ω_{SL}) ranging from 150 to 3,500 Hz, with the absolute offset frequencies (Ω 2π⁻¹, Hz) ranging from 0–3.5× the applied spinlock power (Supplementary Table 1). Magnetization of the spins of interest were allowed to relax under an applied spinlock for the following durations: 0–60 ms for C1'/C6/C8 of hp-GT DNA and C1' of Dickerson-GT, 0–50 ms for C6/C8 Dickerson-GT DNA.

¹³C/¹⁵N CSs of ionized dNTPs and rNTPs. Carbon and nitrogen CSs of neutral and deprotonated dNTPs and rNTPs were collected at 25 °C and pH ≈ 6.9 and ≈ 12.5 using a 2D [¹⁵N, ¹³C] HMQC experiment on a 14.1 T Agilent spectrometer equipped with a Bruker HNCP cryoprobe.

Analysis of R_{1ρ} data. Fitting of ¹³C and ¹⁵N R_{1ρ} data. R_{1ρ} values were calculated by fitting the decay of peak intensity versus relaxation delay to a monoexponential⁵⁸. Errors in R_{1ρ} were estimated using spectral noise and duplicate R_{1ρ} data points using a suite of Mathematica notebooks⁵⁸. Measured on- and off-resonance R_{1ρ} data were globally fit to algebraic equations describing N-site chemical exchange using a Levenberg–Marquardt method, weighted to the experimental error in the R_{1ρ} data. For two-state exchange, data was fit to the Laguerre equation (1)⁵⁹ under the valid assumption that the populations are highly asymmetric, such that p_{GS} ≫ p_{ES} where p_{ES} < 0.3. For the three-state chemical exchange model where k_{BC} = k_{CB} = 0, data was fit to both the three-state Laguerre equation (2) as well as the general three-state equation (3)⁵⁹, where p_{GS} ≫ p_{ES} and p_{ES} < 0.1. Fitted parameters derived from equation (2) and equation (3) are in excellent agreement with one another (Supplementary Table 1). Statistical tests, F-test and Akaike information criterion (AIC), were used to select the best-fit N-state exchange model^{24,25} (Extended Data Fig. 3). ¹³C and ¹⁵N RD data from the dG•dT or rG•rU mispair resonances for each construct, at each temperature and pH condition were fitted globally (where k_{ex} and p_B are the shared-parameters) when possible. In the cases where ¹⁵N RD data was available but ¹³C RD data showed no chemical exchange, did not exhibit three-state exchange, or was not collected, the ¹⁵N N1/N3 RD data were globally fitted as described above. Bloch–McConnell⁶⁰ (B-M) numerical simulations were used to validate the algebraic approximations for two- and three-state exchange scenarios. Two-state Laguerre equation^{59,61}.

$$R_{1\rho} = R_1 \cos^2 \theta + R_2 \sin^2 \theta + \frac{\sin^2 \theta p_{GS} p_{ES1} \Delta \omega_{ES1}^2 k_{ex1}}{\left\{ \omega_{GS}^2 \omega_{ES1}^2 / \omega_{eff}^2 + k_{ex1}^2 - \sin^2 \theta p_{GS} p_{ES1} \Delta \omega_{ES1}^2 \left(1 + \frac{2k_{ex1}^2 (p_{GS} \omega_{GS}^2 + p_{ES1} \omega_{ES1}^2)}{\omega_{GS}^2 \omega_{ES1}^2 + \omega_{eff}^2 k_{ex1}^2} \right) \right\}} \quad (1)$$

Three-state Laguerre equation with no minor exchange²⁵:

$$R_{1\rho} = R_1 \cos^2 \theta + R_2 \sin^2 \theta + \sin^2 \theta \left(\frac{p_{GS} p_{ES1} \Delta \omega_{ES1}^2 k_{ex1}}{\left\{ \omega_{GS}^2 \omega_{ES1}^2 / \omega_{eff}^2 + k_{ex1}^2 - \sin^2 \theta p_{GS} p_{ES1} \Delta \omega_{ES1}^2 \left(1 + \frac{2k_{ex1}^2 (p_{GS} \omega_{GS}^2 + p_{ES1} \omega_{ES1}^2)}{\omega_{GS}^2 \omega_{ES1}^2 + \omega_{eff}^2 k_{ex1}^2} \right) \right\}} \right. \\ \left. + \frac{p_{GS} p_{ES2} \Delta \omega_{ES2}^2 k_{ex2}}{\left\{ \omega_{GS}^2 \omega_{ES2}^2 / \omega_{eff}^2 + k_{ex2}^2 - \sin^2 \theta p_{GS} p_{ES2} \Delta \omega_{ES2}^2 \left(1 + \frac{2k_{ex2}^2 (p_{GS} \omega_{GS}^2 + p_{ES2} \omega_{ES2}^2)}{\omega_{GS}^2 \omega_{ES2}^2 + \omega_{eff}^2 k_{ex2}^2} \right) \right\}} \right) \quad (2)$$

Three-state general equation with no minor exchange^{59,62}:

$$R_{1\rho} = R_1 \cos^2 \theta + R_2 \sin^2 \theta + \sin^2 \theta \left(\frac{k_{GS \rightarrow ES1} \Delta \omega_{ES1}^2}{\Omega_{ES1}^2 + \omega_{SL}^2 + k_{ES1 \rightarrow GS}^2} + \frac{k_{GS \rightarrow ES2} \Delta \omega_{ES2}^2}{\Omega_{ES2}^2 + \omega_{SL}^2 + k_{ES2 \rightarrow GS}^2} \right) \quad (3)$$

in which R_1 and R_2 are the intrinsic longitudinal and transverse relaxation rates (s^{-1}). The exchange rates are defined as $k_{\text{exi}} = k_{\text{GS} \rightarrow \text{ESi}} + k_{\text{ESi} \rightarrow \text{GS}}$, where $k_{\text{GS} \rightarrow \text{ESi}} = p_{\text{ESi}} k_{\text{exi}}$ and $k_{\text{ESi} \rightarrow \text{GS}} = p_{\text{GS}} k_{\text{exi}}$ and where $i = 1$ or 2 . The CS difference between the GS and ESs is given by $\Delta\omega_{\text{ESi}} = \Omega_{\text{ESi}} - \Omega_{\text{GS}}$, where $\Omega = \Omega_{\text{obs}} - \omega_{\text{rf}}$ defines the resonance offset from the carrier frequency (ω_{rf}), $\Omega_{\text{obs}} = (\Omega_{\text{GS}} p_{\text{GS}} + \Omega_{\text{ESi}} p_{\text{ESi}})$ or $\Omega_{\text{obs}} = (\Omega_{\text{GS}} p_{\text{GS}} + \Omega_{\text{ES1}} p_{\text{ES1}} + \Omega_{\text{ES2}} p_{\text{ES2}})$, and where $\Omega_{\text{obs}} \approx \Omega_{\text{GS}}$ when $p_{\text{GS}} \gg p_{\text{ESi}}$, as is the case in the ESs measured. The average effective spinlock field in the rotating frame is given by $\omega_{\text{eff}}^2 = \Omega^2 + \omega_{\text{SL}}^2$ and $\omega_{\text{GS}}^2 = (\Omega_{\text{GS}} - \omega_{\text{rf}})^2 + \omega_{\text{SL}}^2$, $\omega_{\text{ESi}}^2 = (\Omega_{\text{ESi}} - \omega_{\text{rf}})^2 + \omega_{\text{SL}}^2$ and ω_{SL} is the spinlock power. The tilt angle in the rotating frame is given by $\theta = \arctan(\omega_{\text{SL}}/\Omega)$.

Analysis of the three-state exchange model. We repeated measurements of RD profiles for hp-GT dG-N1 and dT-N3 at pH 8.4 and 25 °C on a different spectrometer and also obtained data that is consistent with three-state exchange (Extended Data Fig. 2 and Supplementary Table 1). In addition, we collected one additional spinlock power (800 Hz) for both resonances, and find that the inclusion of this extra spinlock power has negligible effect on the fitted exchange parameters indicating that they are robustly determined by the measured data. We note that differences in the RD profiles and fitted parameters between spectrometers are largely within error, with minor differences likely arising due to small differences in temperature and/or spinlock calibrations.

The dT-N3 resonance of hp-GT DNA at 25 °C and pH 7.9 was also individually refit assuming both two-state and three-state exchange. The individual dT-N3 three-state fit gave very similar fitted parameters as the dT-N3 in the three-state global fit with dG-N1. The F-test (at 0.05 significance level) favoured the three-state individual fit model over the two-state individual fit model. AIC also favoured the three-state model, estimating the three-state model to be 3.1×10^{15} times more likely to be the correct model than the two-state model. The individual two-state and three-state fits to dG-N1 RD profiles give similar agreement ($\chi^2 = 0.83$) when the ES2 CS is fixed based on the globally fitted value; however, statistical tests slightly favour the simpler model (AIC favours two-state by approximately twofold).

The three-state model is supported by statistical tests (F-test and AIC favour the three-state over two-state exchange model) and also by B-M simulations (data not shown) for the hp-GT dG-N1 and dT-N3 data at pH 7.9. Here, R_{1p} values were simulated, noise corrupted, and evaluated for the ability to report on the three-state exchange following the procedure reported in Bothe *et al.*⁶³

Bloch–McConnell three-state numerical simulations. Parameters derived from the algebraic three-state fits, along with the ω_{rf} and ω_{SL} , were used to simulate numerical solutions to the three-state B-M equations⁶⁰ (Extended Data Fig. 3). The B-M simulations were carried out using a procedure similar to that described recently for two-state exchange⁶³. Simulations were carried out assuming a 0.25 s relaxation delay under the applied ω_{SL} .

Thermodynamic analysis of R_{1p} RD-derived parameters. Temperature-dependent analysis of forward and reverse exchange rates in the site-labelled (dG15•dT5 $^{13}\text{C}/^{15}\text{N}$) hp-GT DNA and G/U labelled hp-GU-20 RNA samples were carried out as previously described²⁴. This analysis employed the ES populations and exchange rates obtained based on two-state global fitting of dG15-N1 and dT5-N3 RD data measured in hp-GT DNA at 10, 20, 25 and 30 °C. For hp-GU-20 RNA, the analysis employed populations and exchange rates obtained based on two-state and three-state global fitting of rG16-N1 and rU5-N3 RD data measured at 10, 20, 25 and 30 °C at pH 6.9. Errors in the fitted thermodynamic parameters are given by weighted fits of the modified van't Hoff equation²⁴ to the RD-derived forward and reverse exchange rates and their errors. The NMR sample temperature was calibrated using 99.8% methanol- d_4 (Cambridge Isotope Laboratories) using the equation $T = -16.7467(\Delta\delta)^2 - 52.5130(\Delta\delta) + 419.1381$, where $\Delta\delta$ is the difference in CS (p.p.m.) between the hydroxyl and methyl proton⁶⁴.

Density functional theory geometry optimizations and CS calculations. All DFT calculations³³ were performed using Gaussian 09c (Gaussian, Inc.)⁶⁵ and carried out on the University of Michigan's Advanced Research Computing HPC cluster, Flux, as previously described²⁴ with minor modifications to DFT method and basis set used. Geometry optimizations were carried out using the empirical exchange-correlation functional, M06-2X⁶⁶, with the 6-31+G(d,p) basis-set. The ^{13}C and ^{15}N isotropic magnetic shielding ($\sigma_{13\text{C}}$ and $\sigma_{15\text{N}}$) were calculated using the GIAO method with M06-2X/6-31+G(d,p). CSs of the nucleobases ($\delta_{13\text{C}}$ and $\delta_{15\text{N}}$) were calculated by $\delta_{13\text{C}/15\text{N}} = \sigma_{13\text{C}/15\text{N}} - \sigma_{\text{TMS}/\text{NH}_3}$, where σ_{TMS} and σ_{NH_3} are the isotropic magnetic shieldings calculated for the reference compounds trimethylsilane (^{13}C) and NH_3 (^{15}N), respectively.

Input structures for DFT calculations. We generated idealized B-/A-form helices corresponding to our sequence contexts (hp-GT DNA and hp-GU-24 RNA) using make-NA (J. Stroud, make-NA, <http://structure.usc.edu/make-na/server.html>; 2011). The duplexes were truncated to the trinucleotide step centred on the target mispair (GTG/CGC for hp-GT DNA and GUA/UGC for hp-GU-24 RNA). The sugar and phosphate moieties were removed and replaced with a methyl for i (dG•dT or rG•rU mispair), $i + 1$, and $i - 1$ base pairs to save on computational

time. Although the structures lack a 2'-deoxyribose or ribose sugar moieties, they will be denoted as dG•dT or rG•rU to avoid confusion. All heavy atoms were frozen for the $i + 1$ and $i - 1$ base pairs while geometry optimizations were carried out for protons and heavy atoms of the central dG•dT or rG•rU base pair as well as the protons of the $i + 1/i - 1$ pairs. We performed full geometry optimizations on: dG•dT and rG•rU GS WB pairs, dG^{enol}•dT, dG•dT^{enol}, rG^{enol}•rU, rG•rU^{enol}, and dG^{enol}•dT and rG^{enol}•rU ES WC-like pairs. All converged to the expected WB or WC-like geometries. In the instance where the geometry optimizations of dG•dT^{enol} and rG•rU^{enol} starting states failed to converge to a stable WC-like dG•dT^{enol}/rG•rU^{enol} geometry (and instead converged to an iWB geometry, in vacuum), the WC-like dG•dT^{enol}/rG•rU^{enol} states were converted to dG•dT^{enol}/rG•rU^{enol} and geometry optimizations were carried out on the protons only. CSs for each state were calculated and later used in calculating population weighted CSs assuming different populations of these mispair species. In addition, while the sugar moieties were truncated to methyl groups to save on computation time, it should be noted that previous DFT studies of nucleotides have shown that tautomerization, primarily of pyrimidines, can have an effect on the sugar conformation⁶⁷. However, we can rule out large changes in sugar pucker arising in ES1 based on the negligibly small chemical exchange contributions to both dG-C1' and dT-C1' (see Extended Data Fig. 2) and only very small changes in dG-C1' and dT-C1' CSs upon locking the enol-like form with $^{16}\text{dG}\cdot\text{dT}$ ($|\Delta\omega_{\text{C1'}}| \leq 0.7$ p.p.m., see Fig. 3b and Extended Data Fig. 7). Finally, while the CSs of the anionic pairs are predicted assuming planar pair geometry, prior computational studies of G•T[−] and G[−]•T nucleobase pairs in isolation have shown that they can favour non-planar and non-WC geometry³⁵.

dG^{enol}•dT distance dependent DFT calculations. We carried out distance dependent DFT calculations⁶⁸ on a pair of WC-like dG^{enol}•dT N1/N9-methyl nucleobases in vacuum using the M06-2X method and 6-31+G(d,p) basis-set, as described above. The geometry of a dG•dT wobble pair and dG^{enol}•dT WC-like pair was optimized with no constraints before CS calculations. The ideal N1-N3 distance of the dG^{enol}•dT WC-like mispair was then manually varied from 2.44 Å to 3.8 Å in increments of 0.1 Å from 2.44–3.04 Å and then to 3.8 Å (Extended Data Fig. 4). At every increment the proton positions alone were optimized and CSs were calculated relative to an optimized dG•dT wobble base pair.

Population-weighted average DFT-predicted CS calculations. Based on the computationally predicted energetic differences between interconverting dG^{enol}•dT and dG•dT^{enol} base mispairs in water ($\Delta G = 0.7$ – 0.8 kcal mol^{−1})³¹ and in a weakly polar medium ($\Delta G = 0.99$ kcal mol^{−1})⁶⁹, we can predict that the dG•dT ES1 CSs represent a population weighted average between interconverting dG^{enol}•dT (~80%) \rightleftharpoons dG•dT^{enol} (~20%) states. Thus, the DFT-predicted CSs for dG^{enol}•dT and dG•dT^{enol} were summed in a population-weighted manner. It is noted that the computationally predicted energetic stabilities of the tautomeric states differ when calculated in water versus vacuum, or a weakly polar medium, with the values predicted in water giving the greatest agreement with our experimental results.

In the case of rG•rU ES1, computational studies have shown that dG•dU^{enol} is ~ 1 kcal mol^{−1} more stable than dG•dT^{enol} in a DNA fragment³⁶, suggesting that an rG^{enol}•rU \rightleftharpoons rG•rU^{enol} equilibrium should be tilted slightly more towards rG•rU^{enol} than dG•dT^{enol} in DNA. We can qualitatively estimate the relative stability between rG^{enol}•rU and rG•rU^{enol} to be 60:40 based on a best fit to the RD-derived CSs.

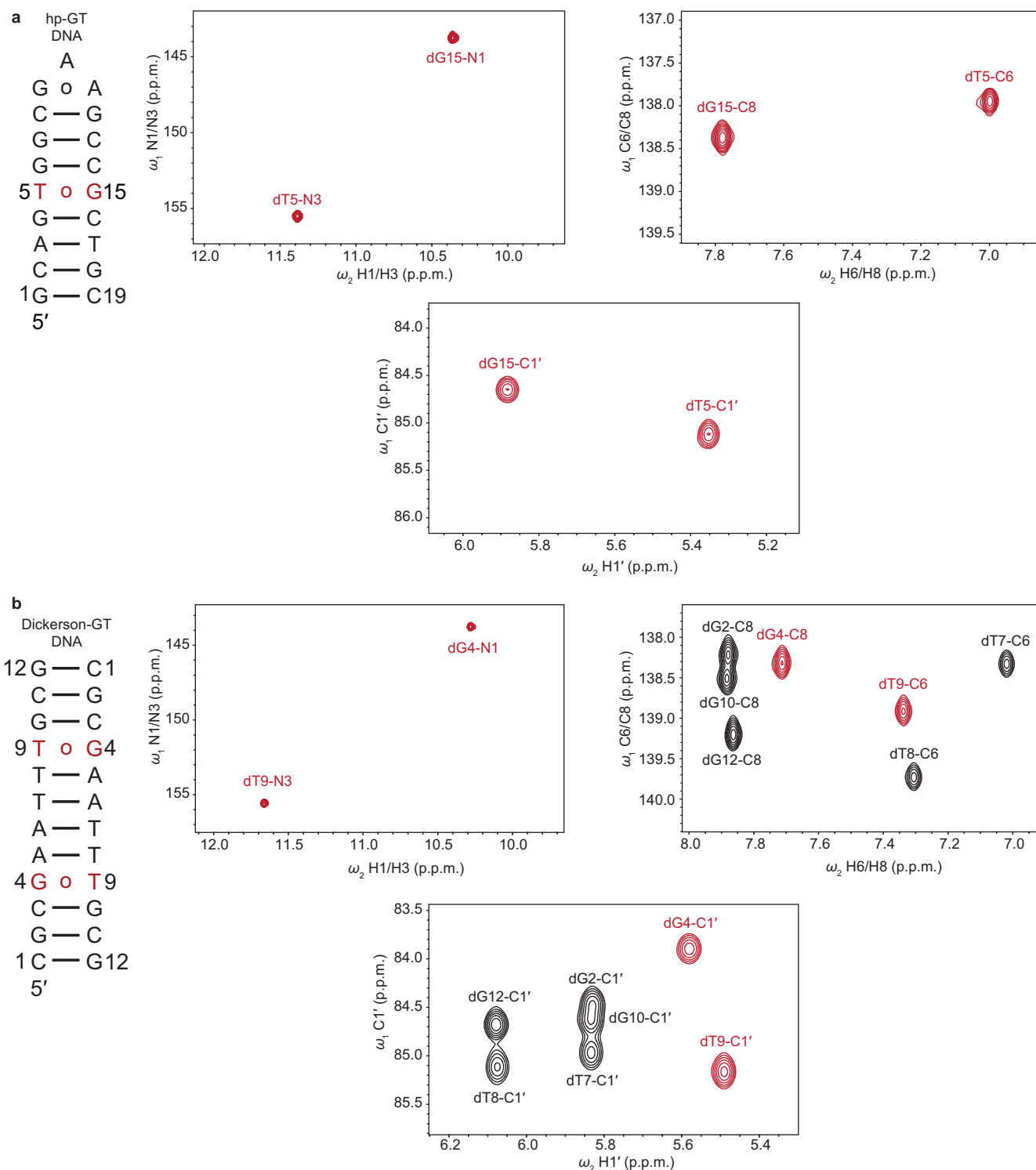
pK_a fitting and probability estimation. The apparent pK_as for hp-GT ES2, ^{58}BrU -hp-GT ES2, and dGTP•dT misincorporation (pH 6.5–8.6) were fit to the Henderson–Hasselbalch equation using a Monte-Carlo (MC) approach. Here, $10^6 p_{\text{B}}$ s at pH 7.9 and/or 8.4 were selected from a Gaussian distribution with mean p_{B} value and standard deviation representing the uncertainty in p_{B} based on fitting of the RD data. 10^6 fits to equation (4) were then carried out assuming these p_{B} values to generate 10^6 pK_as.

$$p_{\text{B}} = \left(\frac{10^{\text{pH} - \text{pK}_a}}{1 + 10^{\text{pH} - \text{pK}_a}} \right) \quad (4)$$

Where p_{B} is the probability of forming ES2 or dGTP•dT misincorporation probability at a given pH. The resulting fitted pK_a values were fitted to a Gaussian distribution. The mean value of the Gaussian distribution is the reported pK_a value and the standard deviation is assumed to be the error. An analogous approach was used to back-calculate predicted p_{B} s at a given pH using the pK_a derived by the above method.

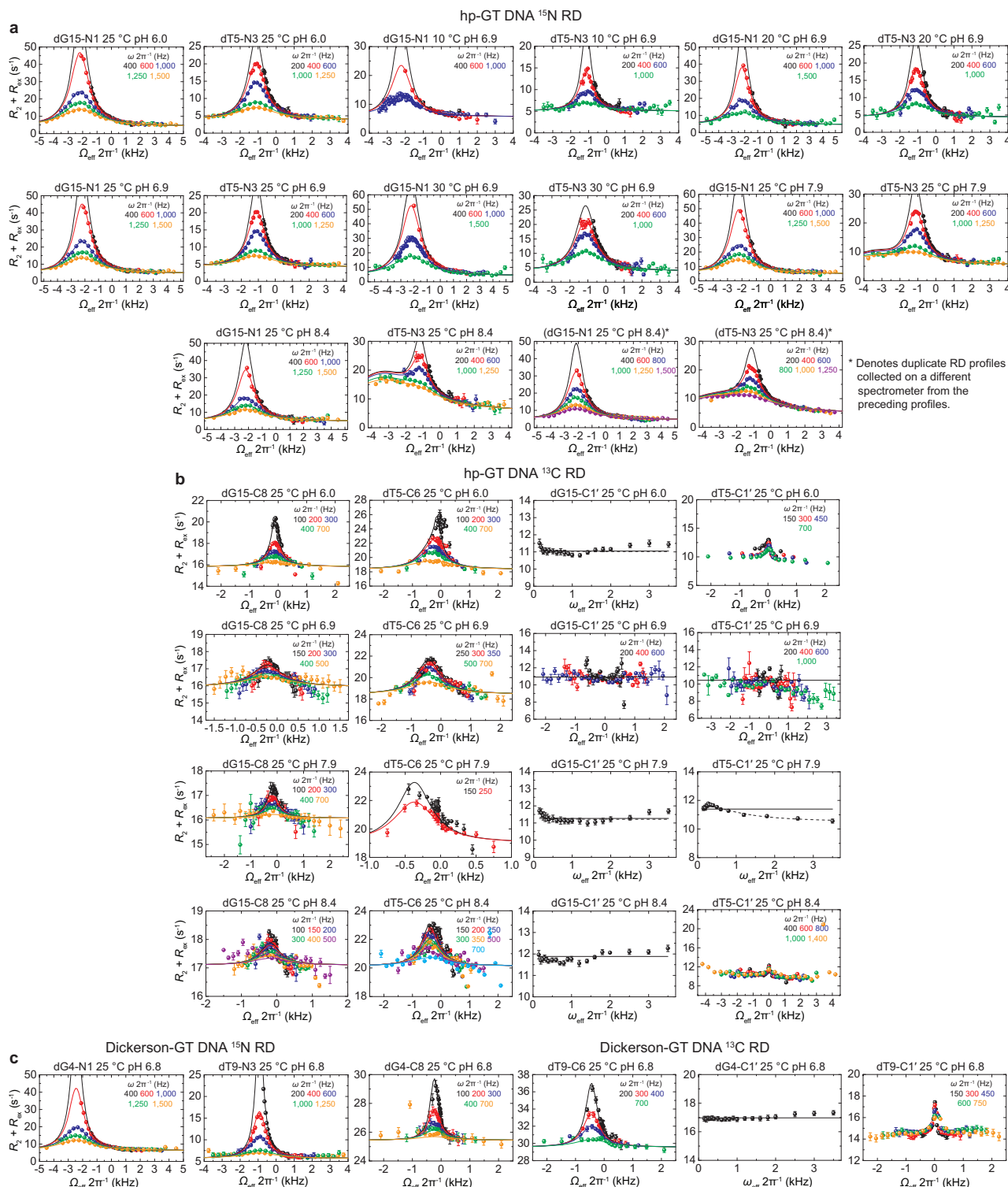
- Zimmer, D. P. & Crothers, D. M. NMR of enzymatically synthesized uniformly $^{13}\text{C}/^{15}\text{N}$ -labeled dna oligonucleotides. *Proc. Natl Acad. Sci. USA* **92**, 3091–3095 (1995).
- Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
- Cavanagh, J., Fairbrother, W. J., Palmer, A. G. III & Skelton, N. J. *Protein NMR spectroscopy: principles and practice* (Academic Press, 1995).

54. Wijmenga, S. S. & van Buuren, B. N. M. The use of NMR methods for conformational studies of nucleic acids. *Prog. Nucl. Magn. Reson. Spectrosc.* **32**, 287–387 (1998).
55. Farjon, J. *et al.* Longitudinal-relaxation-enhanced NMR experiments for the study of nucleic acids in solution. *J. Am. Chem. Soc.* **131**, 8571–8577 (2009).
56. Sathyamoorthy, B., Lee, J., Kimsey, I., Ganser, L. & Al-Hashimi, H. Development and application of aromatic [^{13}C , ^1H] SOFAST-HMQC NMR experiment for nucleic acids. *J. Biomol. NMR* **60**, 77–83 (2014).
57. Nikolova, E. N., Gottardo, F. L. & Al-Hashimi, H. M. Probing transient Hoogsteen hydrogen bonds in canonical duplex DNA using NMR relaxation dispersion and single-atom substitution. *J. Am. Chem. Soc.* **134**, 3667–3670 (2012).
58. Spyropoulos, L. A suite of Mathematica notebooks for the analysis of protein main chain ^{15}N NMR relaxation data. *J. Biomol. NMR* **36**, 215–224 (2006).
59. Palmer, A. G. III & Massi, F. Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev.* **106**, 1700–1719 (2006).
60. McConnell, H. M. Reaction rates by nuclear magnetic resonance. *J. Chem. Phys.* **28**, 430–431 (1958).
61. Miloushev, V. Z. & Palmer, A. G. III. $R_{1\rho}$ relaxation for two-site chemical exchange: general approximations and some exact solutions. *J. Magn. Reson.* **177**, 221–227 (2005).
62. Trott, O. & Palmer, A. G. III. Theoretical study of $R_{1\rho}$ rotating-frame and R_2 free-precession relaxation in the presence of n -site chemical exchange. *J. Magn. Reson.* **170**, 104–112 (2004).
63. Bothe, J. R., Stein, Z. W. & Al-Hashimi, H. M. Evaluating the uncertainty in exchange parameters determined from off-resonance $R_{1\rho}$ relaxation dispersion for systems in fast exchange. *J. Magn. Reson.* **244**, 18–29 (2014).
64. Findeisen, M., Brand, T. & Berger, S. A ^1H -NMR thermometer suitable for cryoprobes. *Magn. Reson. Chem.* **45**, 175–178 (2007).
65. Gaussian 09 (Gaussian, Inc., Wallingford, CT, USA, 2009).
66. Zhao, Y. & Truhlar, D. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **120**, 215–241 (2008).
67. Kochina, O. S., Zhuravitsky, R. O. & Hovorun, D. M. Effect of the nucleotide bases tautomerization on the conformational properties of the nucleosides: quantum-mechanical investigation by the DFT method. *Rep. Natl. Acad. Sci. Ukraine* **1**, 181–186 (2008).
68. Czernek, J., Fiala, R. & Sklenář, V. r. Hydrogen bonding effects on the ^{15}N and ^1H shielding tensors in nucleic acid base pairs. *J. Magn. Reson.* **145**, 142–146 (2000).
69. Brovarets', O. O. & Hovorun, D. M. The nature of the transition mismatches with Watson–Crick architecture: the G^*T or G^*T^* DNA base mispair or both? A QM/QTAIM perspective for the biological problem. *J. Biomol. Struct. Dyn.* **33**, 925–945 (2014).
70. Brovarets', O. O. & Hovorun, D. M. Physicochemical mechanism of the wobble DNA base pairs Gua^*Thy and Ade^*Cyt transition into the mismatched base pairs Gua^*Thy and Ade^*Cyt^* formed by the mutagenic tautomers. *Ukr. Bioorg. Acta* **8**, 12–18 (2009).
71. Serganov, A. *et al.* Structural basis for discriminative regulation of gene expression by adenine- and guanine-sensing mRNAs. *Chem. Biol.* **11**, 1729–1741 (2004).
72. Boosalis, M. S., Petruska, J. & Goodman, M. F. DNA polymerase insertion fidelity. Gel assay for site-specific kinetics. *J. Biol. Chem.* **262**, 14689–14696 (1987).
73. Mendelman, L. V., Boosalis, M. S., Petruska, J. & Goodman, M. F. Nearest neighbor influences on DNA polymerase insertion fidelity. *J. Biol. Chem.* **264**, 14415–14423 (1989).
74. Singer, B., Chavez, F., Goodman, M. F., Essigmann, J. M. & Dosanjh, M. K. Effect of 3' flanking neighbors on kinetics of pairing of dCTP or dTTP opposite O^6 -methylguanine in a defined primed oligonucleotide when *Escherichia coli* DNA polymerase I is used. *Proc. Natl. Acad. Sci. USA* **86**, 8271–8274 (1989).
75. Dosanjh, M. K., Essigmann, J. M., Goodman, M. F. & Singer, B. Comparative efficiency of forming $\text{m}^4\text{T}^*\text{G}$ versus $\text{m}^4\text{T}^*\text{A}$ base pairs at a unique site by use of *Escherichia coli* DNA polymerase I (Klenow fragment) and *Drosophila melanogaster* polymerase α -primase complex. *Biochemistry* **29**, 4698–4703 (1990).
76. Singer, B. & Dosanjh, M. K. Site-directed mutagenesis for quantitation of base-base interactions at defined sites. *Mutat. Res. Fund. Mol. Mech. Mut.* **233**, 45–51 (1990).
77. Dosanjh, M. K., Galeros, G., Goodman, M. F. & Singer, B. Kinetics of extension of O^6 -methylguanine paired with cytosine or thymine in defined oligonucleotide sequences. *Biochemistry* **30**, 11595–11599 (1991).



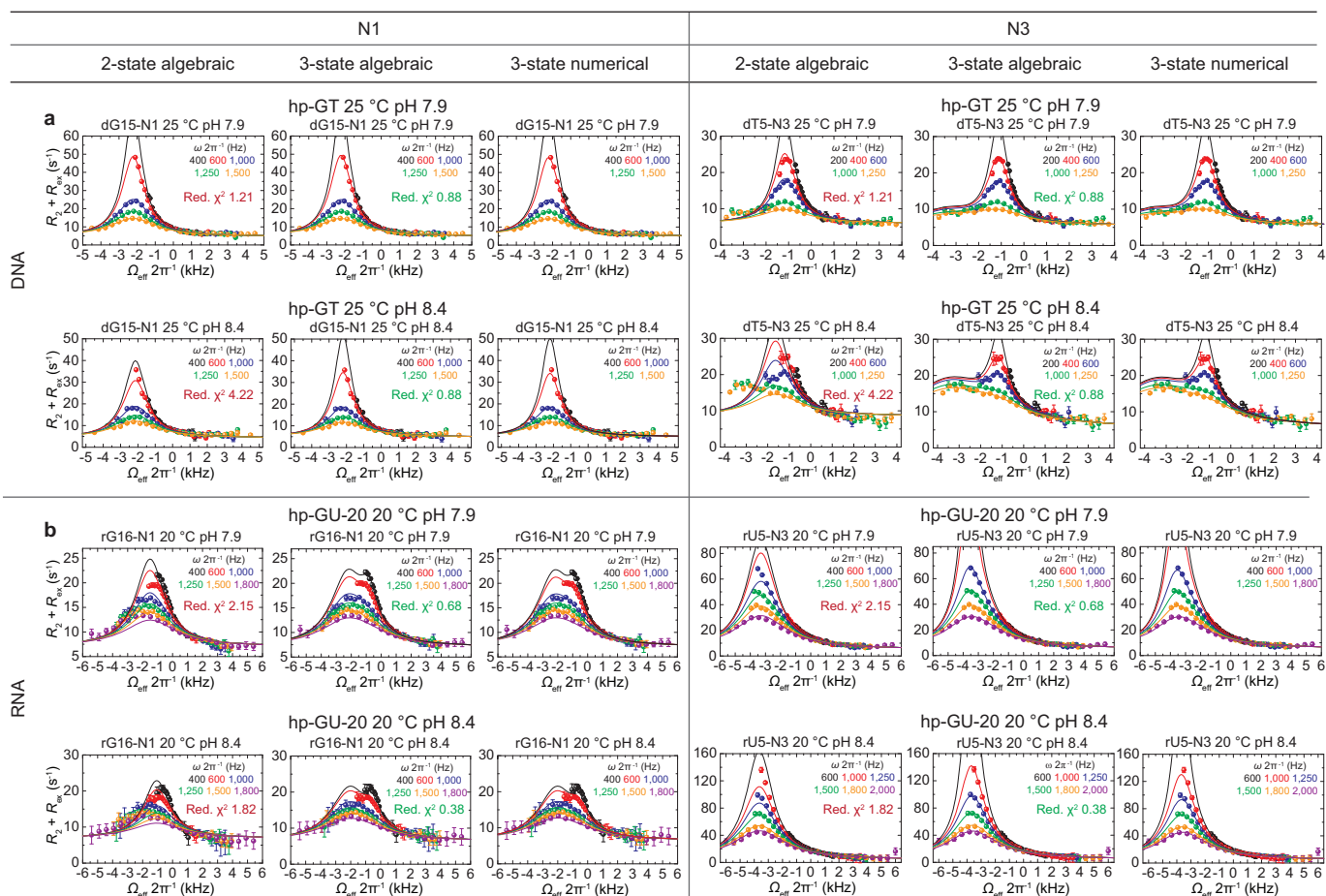
Extended Data Figure 1 | NMR spectra of site- and selectively-labelled dG•dT mispair DNA constructs. **a, b,** Shown are the hp-GT DNA (**a**) and Dickerson-GT (**b**) constructs with $^{13}\text{C}/^{15}\text{N}$ labelled dG•dT mispairs

highlighted in red along with 2D imino [^{15}N , ^1H] HSQC, 2D aromatic [^{13}C , ^1H] HSQC and 2D C1' [^{13}C , ^1H] HSQC spectra (pH 6.9, 25 °C).



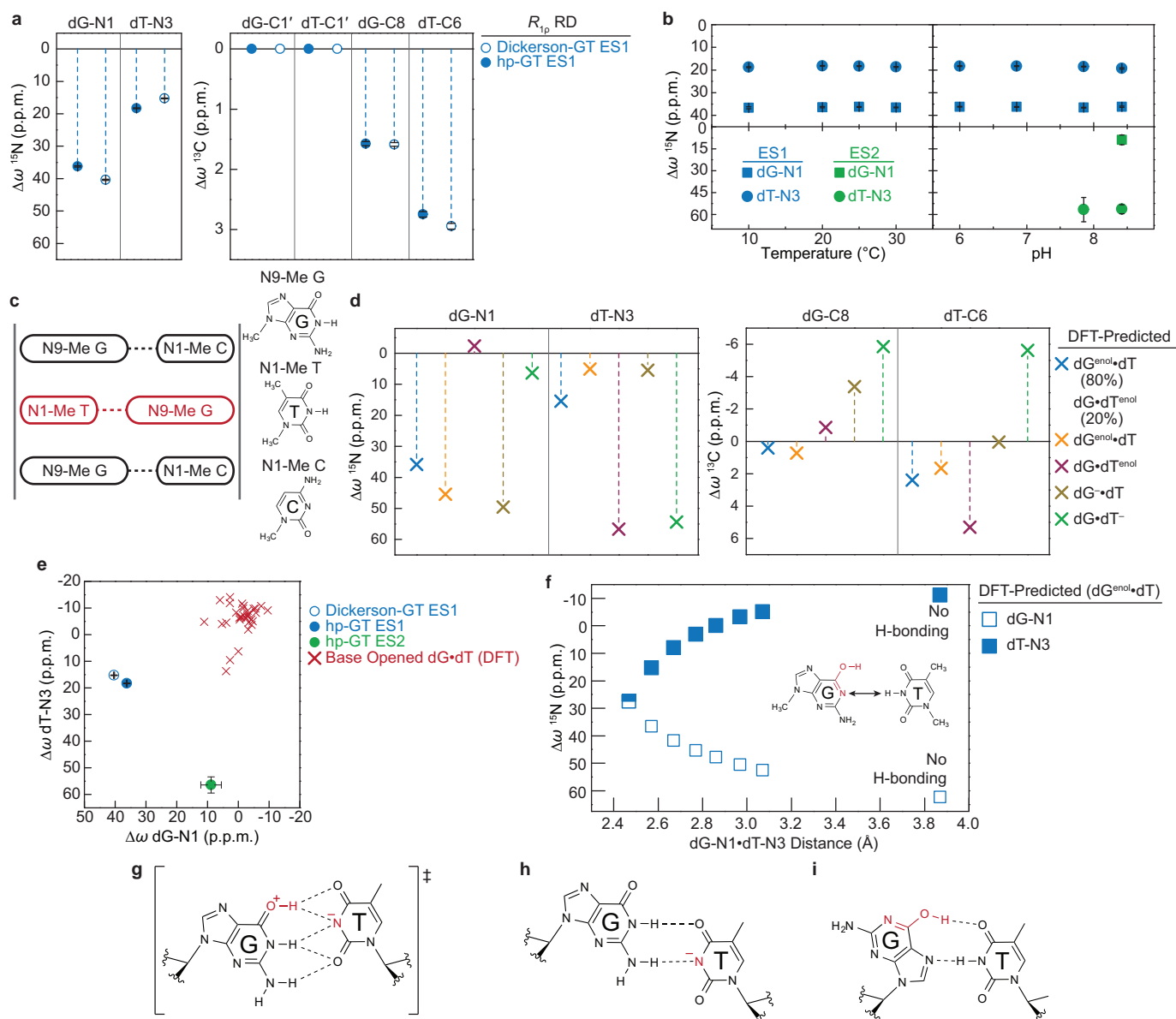
Extended Data Figure 2 | Rotating frame relaxation dispersion profiles of dG•dT mispairs in hp-GT and Dickerson-GT DNA constructs. RD profiles showing chemical exchange ($R_2 + R_{ex}$) in the dG•dT mispair as a function of the spin lock offset ($\Omega_{\text{eff}} 2\pi^{-1}$) and spin lock power ($\omega_{\text{SL}} 2\pi^{-1}$, colour coded in insets). **a**, **b**, Shown are ^{15}N (**a**) and ^{13}C (**b**) RD profiles in hp-GT DNA. On-resonance profiles showing solid and dashed black lines indicate

process (dash). The hp-GT dG15-N1 and dT5-N3 in brackets denote duplicate profiles (with an additional 800 Hz spinlock power for each) collected at pH 8.4 and 25 °C collected on a different spectrometer from the preceding profiles. **c**, ^{15}N and ^{13}C RD profiles for Dickerson-GT. Sample conditions are indicated on each profile. Error bars represent experimental uncertainty (one s.d., see Methods).



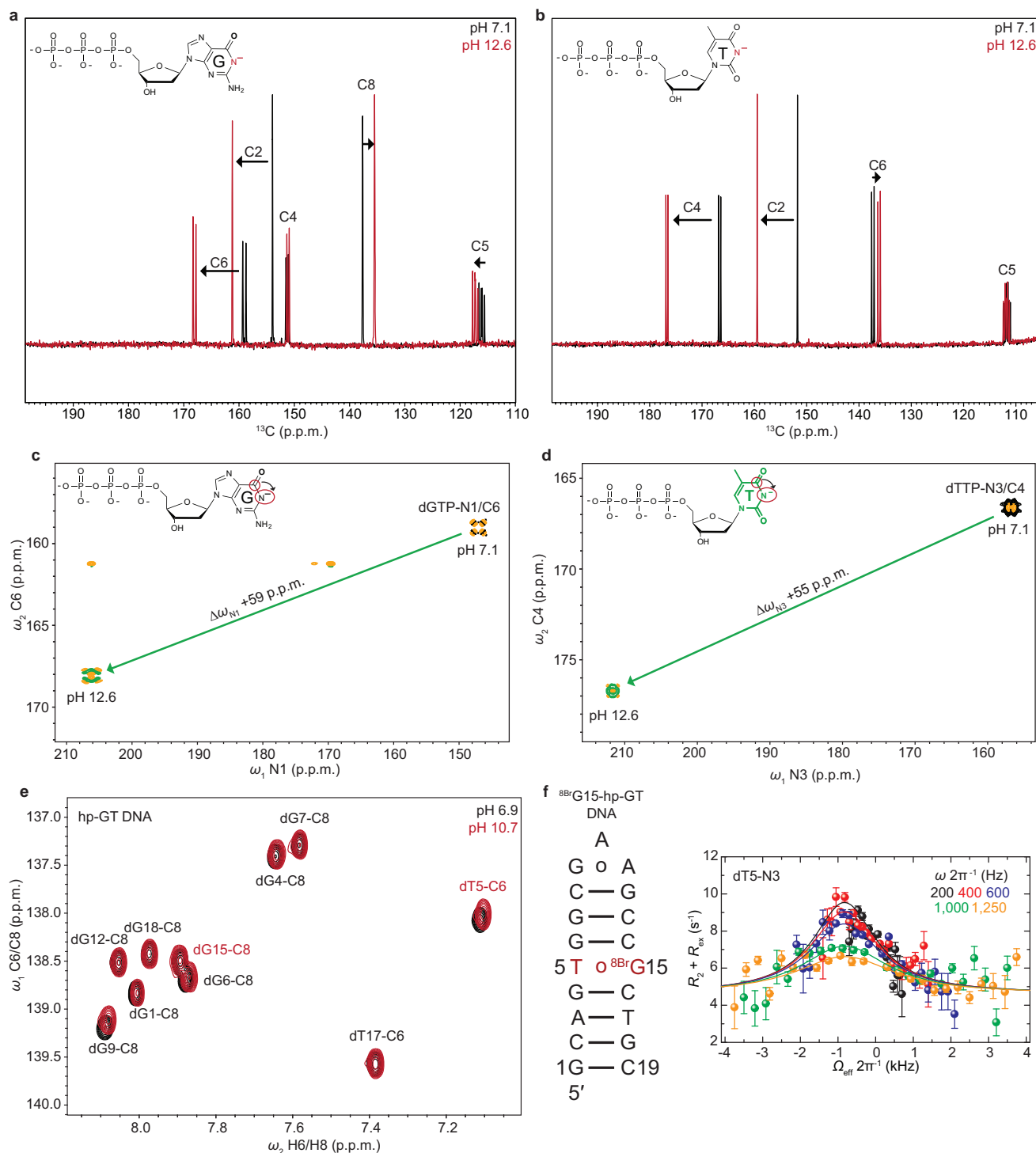
Extended Data Figure 3 | Multiple site exchange comparison and numerical solutions. **a, b,** Global fitting of hp-GT DNA (**a**) and hp-GU-20 RNA (**b**) N1 and N3 RD profiles to two-state algebraic equation (equation (1), fit reduced χ^2 shown in inset) and three-state algebraic equation (equation (2), fit reduced χ^2 shown in inset). Numerical solutions to the Bloch–McConnell three-state

equations assuming no minor exchange and input exchange parameters obtained based on the three-state algebraic fit are also shown to establish the validity of the three-state expression under these exchange scenarios (equation (2), see Methods). Sample conditions are indicated on each profile. Error bars represent experimental uncertainty (one s.d., see Methods).



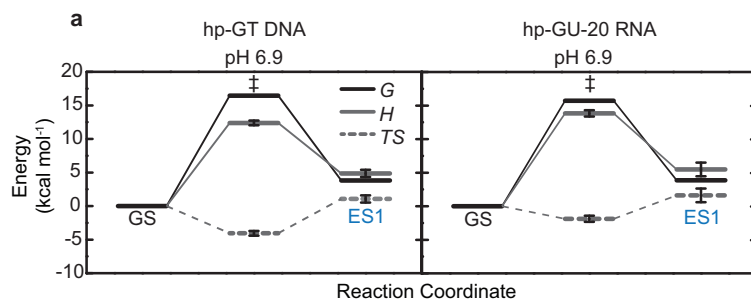
Extended Data Figure 4 | Chemical shift fingerprinting dG•dT excited states. **a**, RD-derived dG15-N1 and dT5-N3 chemical shifts (CSs) (referenced to GS WB) for ES1 (25 °C and pH 6.9) and ES2 (25 °C and pH 8.4) of hp-GT and ES1 of Dickerson-GT (25 °C and pH 6.9) are shown. Errors in all RD-derived fitted parameters (for example, $\Delta\omega$) reflect experimental uncertainty (one s.d.) from the weighted global fit (see Methods). **b**, RD-derived hp-GT dG•dT ES1 (blue) and ES2 (green) ^{15}N CSs are shown as a function of temperature and pH for both dG15-N1 (square) and dT5-N3 (circle). **c**, Scheme used to calculate CSs using DFT (see Methods). Shown is a schematic representation of scenario used for calculating CSs using DFT. Idealized B-form DNA helix is generated to give a central dG•dT mispair (red) that is flanked by canonical dG•dC pairs, analogous to the hp-GT construct. Residues are trimmed to 1-/9-methyl bases and $i + 1/i - 1$ pairs are frozen in place for subsequent geometry optimizations and NMR CS calculations.

d, DFT-calculated CSs (referenced to an energy optimized WB geometry) are shown for various tautomeric and anionic configurations, where dG^{enol}•dT/dG•dT^{enol} represents population-weighted average over dG^{enol}•dT (80%) and dG•dT^{enol} (20%). **e**, RD-derived ES1 and ES2 CSs are plotted against DFT-calculated CSs of base opened dG•dT mispairs, taken from X-ray structures and pruned to 1-/9-methyl bases. **f**, DFT-calculated CSs (referenced to an energy optimized WB geometry) are plotted as a function of dG-N1 to dT-N3 inter-atomic distance for a WC-like dG^{enol}•dT tautomeric pair. **g**, Computational studies^{31,32,70} predict that the tautomeric pathway proceeds via a planar dG⁺•dT⁻ ion pair (charge delocalization is implied) that is highlighted by a network of five H-bonds. **h**, Predicted pair geometry of an anionic dG•dT⁻ inverted wobble. Deprotonated dT-N3 is highlighted in red (charge delocalization is implied). **i**, Predicted pair geometry of a dG^{enol}•dT Hoogsteen mispair.



Extended Data Figure 5 | Attempts to trap anionic dG. **a**, 1D ^{13}C spectra (without ^{13}C - ^{13}C homonuclear decoupling) of the aromatic carbon region of protonated dGTP (black) and anionic dGTP (red) showing CS perturbations induced upon deprotonation of dGTP-N1. **b**, ^{13}C spectra (without ^{13}C - ^{13}C homonuclear decoupling) of the aromatic carbon region of protonated dTTP (black) and anionic dTTP (red) showing CS perturbations induced upon deprotonation of dTTP-N3. **c**, 2D ^{15}N - ^{13}C HMQC spectra of dGTP showing CS of dGTP-N1 induced upon deprotonation. The spectra is rotated by 90° , to depict ^{15}N CS along x-axis for visualization purposes. Red circles on inset structure highlight measured resonances (C6 and N1). **d**, 2D ^{15}N - ^{13}C HMQC spectra of dTTP showing CS perturbation of dTTP-N3 induced upon

deprotonation. The spectra is rotated by 90° , to depict ^{15}N CS along x-axis for visualization purposes. Red circles on inset structure highlight measured resonances (C4 and N3). **e**, hp-GT DNA spectra of the dG/dT aromatic carbons upon increase in pH from 6.9 (black) to 10.7 (red). Minor upfield CSs are observed for dT5-C6 and dG9-C8, but not dG15-C8, indicating the dT5 in the dG•dT mispair is likely undergoing deprotonation and not the paired dG15. **f**, $^{8\text{Br}}\text{G15}$ -hp-GT DNA construct bearing a $^{13}\text{C}/^{15}\text{N}$ site-labelled dT5 paired with a 8-bromo-2'-deoxyguanosine is shown (left) along with the ^{15}N RD profile for the paired dT5-N3. Error bars represent experimental uncertainty (one s.d., see Methods).



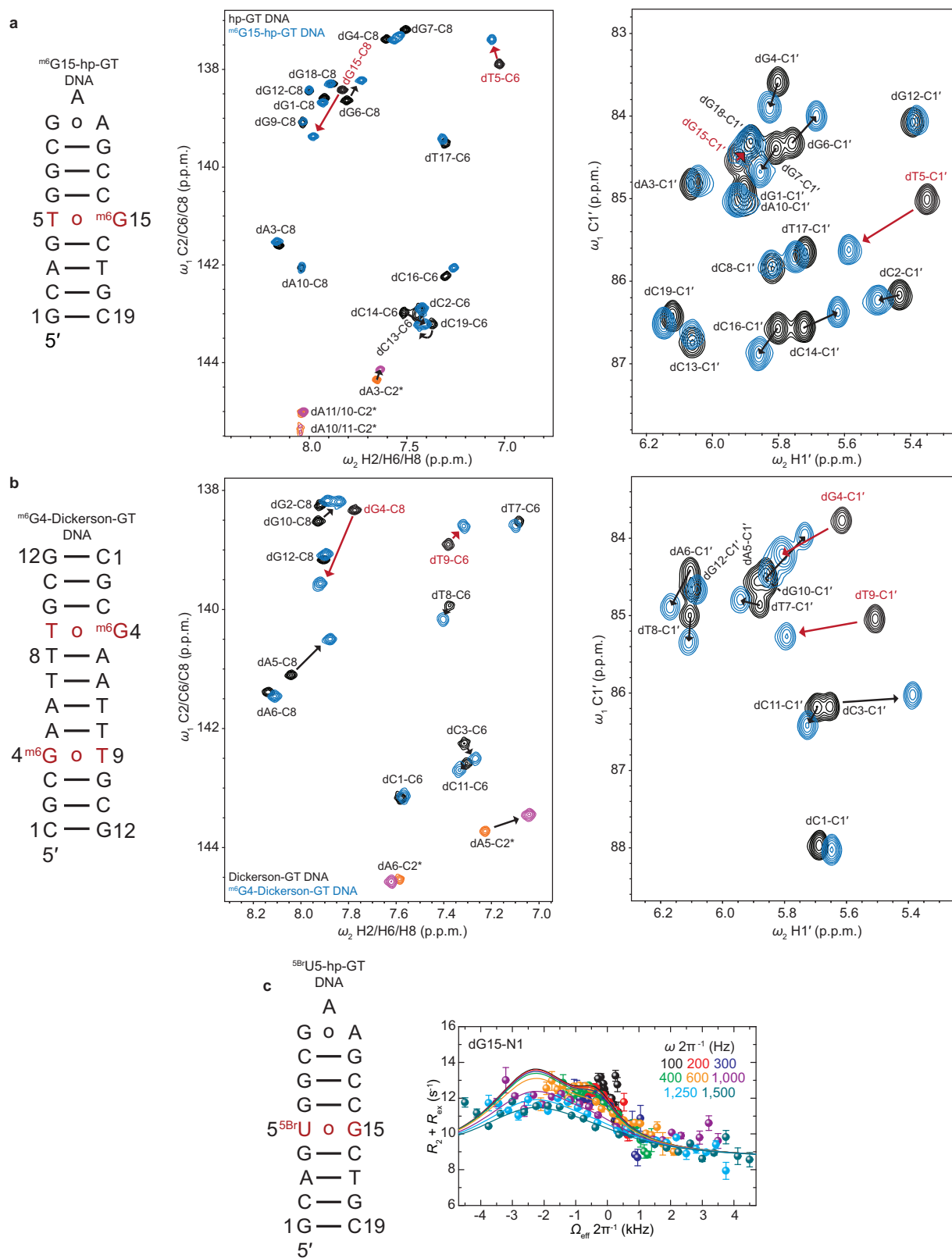
b

Construct	pH	Mispair	ES	ΔG_{AB}	ΔH_{AB}	$T\Delta S_{AB}$	ΔG_A^T	ΔH_A^T	$T\Delta S_A^T$	ΔG_B^T	ΔH_B^T	$T\Delta S_B^T$
hp-GT DNA	6.9	dG15•dT5	ES1	3.80±0.01	4.87±0.53	1.07±0.53	16.4±0.01	12.4±0.35	-4.06±0.35	12.6±0.01	7.51±0.40	-5.13±0.40
hp-GU-20 RNA	6.9	rG16•rU5	ES1	3.86±0.02	5.48±1.03	1.62±1.03	15.7±0.01	13.8±0.46	-1.87±0.46	11.8±0.01	8.36±0.92	-3.48±0.92
* ^{5Br} U-hp-GT DNA	6.9	dG15• ^{5Br} dU5	ES1	4.67±0.30	--	--	16.0±0.21	--	--	11.3±0.21	--	--
* hp-GT DNA	8.4	dG15•dT5	ES2	4.59±0.24	--	--	16.4±0.17	--	--	11.8±0.17	--	--
* hp-GU-20 RNA	8.4	rG16•rU5	ES2	3.29±0.11	--	--	15.9±0.08	--	--	12.6±0.08	--	--
* ^{5Br} U-hp-GT DNA	6.9	dG15• ^{5Br} dU5	ES2	3.37±0.40	--	--	15.1±0.28	--	--	11.8±0.27	--	--

(kcal mol⁻¹)

Extended Data Figure 6 | Kinetic-thermodynamic plots and parameters.
a, Kinetic-thermodynamic diagram for exchange between GS and ES1 via a transition state for hp-GT DNA ES1 (left) and hp-GU-20 RNA ES1 (right), showing activation (G^\ddagger) and net free energy (G), enthalpy (H), and entropy (TS) changes (referenced to 0). **b**, Kinetic-thermodynamic parameters derived from RD data. Asterisk denotes parameters calculated using only a single

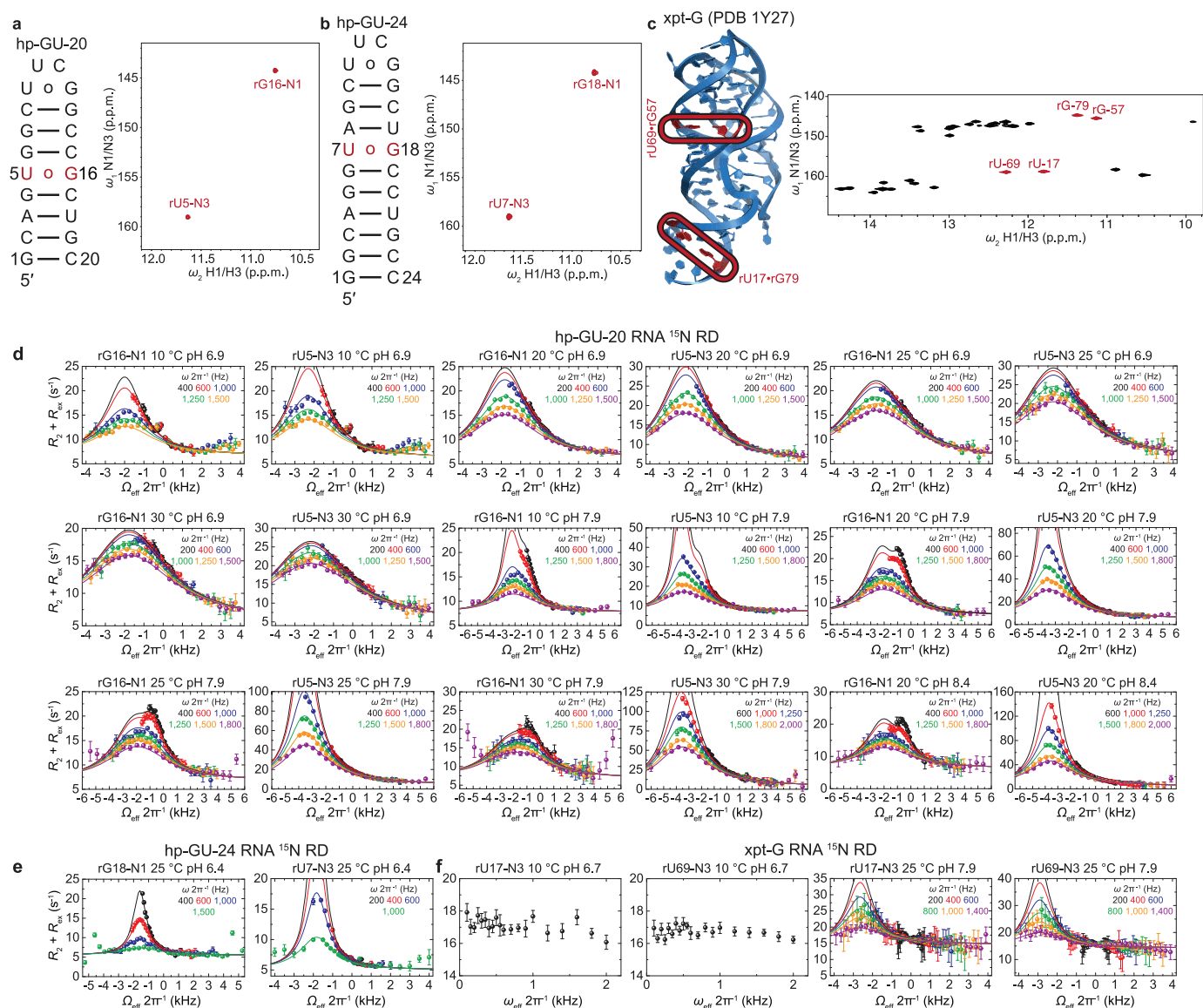
temperature (see Methods), wherein enthalpic and entropic parameters cannot be derived. Here, dG15•dT5 ES2 values were calculated at 25.05 °C, rG16•rU5 ES2 values were calculated at 20.05 °C, and dG15•^{5Br}dU5 ES1 and ES2 values were calculated at 10.05 °C. Error reflects experimental uncertainty (one s.d.) of the weighted global fits of the corresponding RD profiles. Error is propagated using the respective uncertainties in k_{ex} and p_{ES} .



Extended Data Figure 7 | Trapping or stabilizing dG•dT ES1 and ES2.

a. ^{m6}G15-hp-GT DNA construct is shown (left) where dG15 is methylated at the O⁶ position to trap a near-WC “dG^{enol}•dT”-like geometry (Fig. 3c). CS perturbations induced in the aromatic (centre) and sugar (right) resonances upon O⁶-methylation (blue) with the hp-GT DNA spectra (black) with the resonances for the dG•dT mispair from hp-GT DNA in red. ^{m6}dG•dT mispair

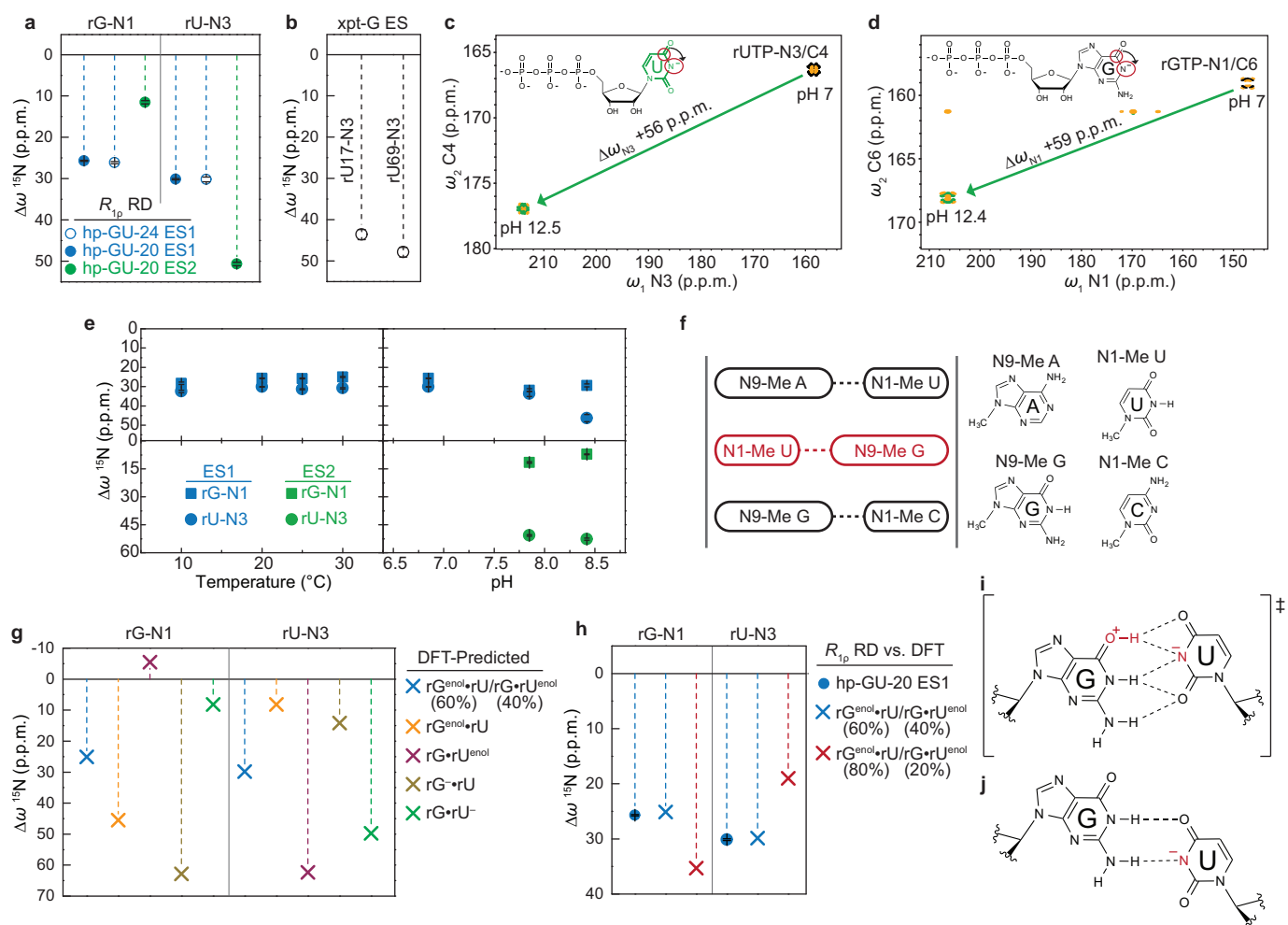
and CSs are highlighted in red. **b.** Similarly, ^{m6}G4-Dickerson-GT DNA construct is shown (left) where dG4 is O⁶-methylated to trap a WC-like state, with similar colour scheme as **a.** **c.** ^{5Br}U5-hp-GT DNA construct bearing a ¹³C/¹⁵N site-labelled dG15 paired with a 5-bromo-2'-deoxythymidine is shown (left) along with the ¹⁵N RD profile for the paired dG15-N1. Error bars represent experimental uncertainty (one s.d., see Methods).



Extended Data Figure 8 | Rotating frame relaxation dispersion profiles for rG•rU mismatches in hp-GU-20, hp-GU-24 and xpt-G RNA constructs.

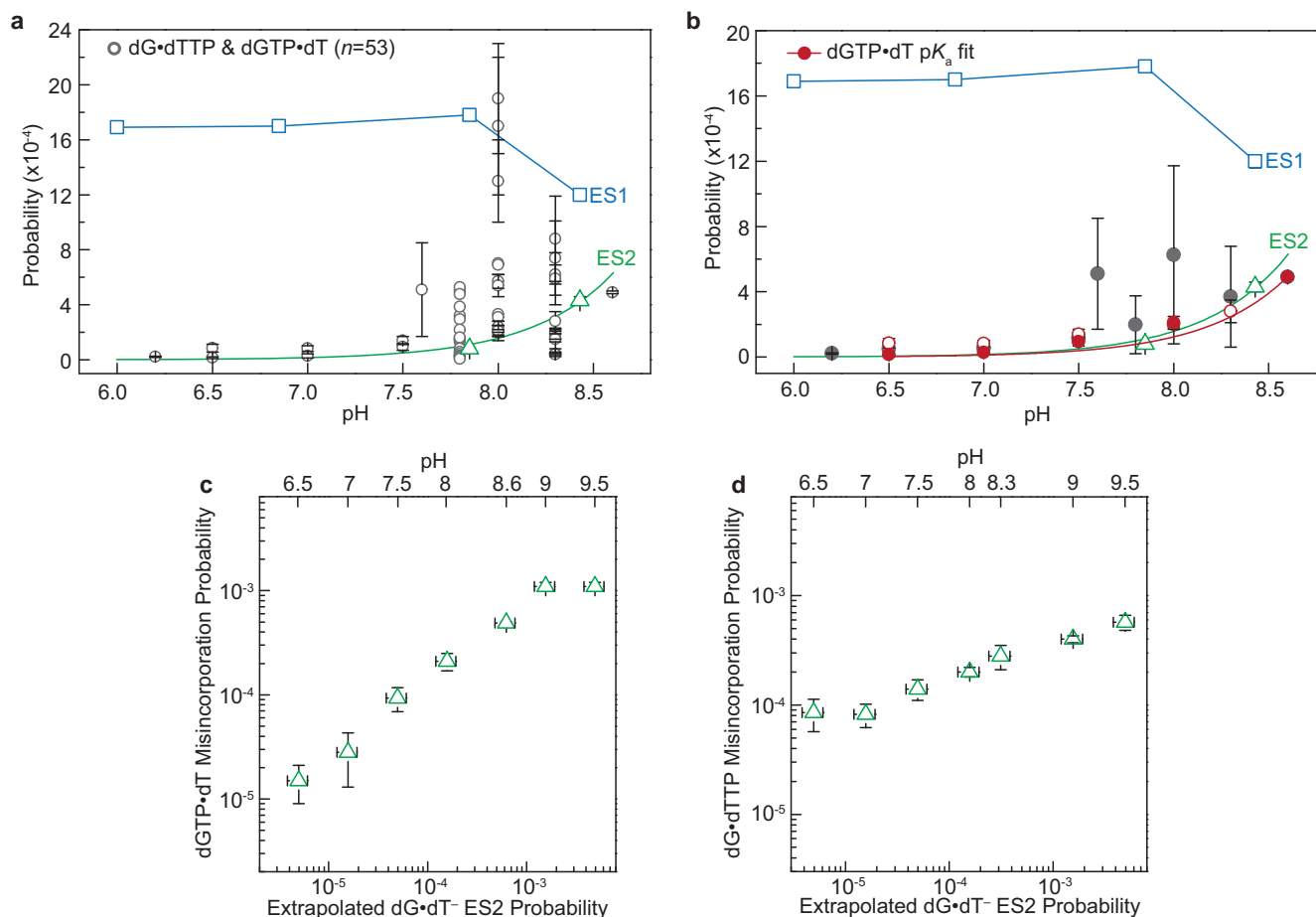
a, b, RNA constructs and the imino [^{15}N , ^1H] HSQC zoomed into the rG•rU wobble region of the spectra for hp-GU-20 and hp-GU-24. rG•rU mismatch resonances are shown in red. **c**, The *Bacillus subtilis* guanine binding riboswitch

(xpt-G RNA)⁷¹ construct and full imino [^{15}N , ^1H] HSQC of folded and guanine ligand-bound riboswitch. rG•rU mismatch resonances are shown in red. **d–f**, ^{15}N RD profiles for hp-GU-20 (**d**) hp-GU-24 (**e**) and xpt-G (**f**) riboswitch RNA. Error bars represent experimental uncertainty (one s.d., see Methods).



Extended Data Figure 9 | CS fingerprinting $rG\cdot rU$ excited states. **a**, RD-derived $rG16\cdot N1$ and $rU5\cdot N3$ CSs (referenced to GS WB) are shown for ES1 (20 °C and pH 6.9) and ES2 of hp-GU-20 (20 °C and pH 7.9) and ES1 $rG18\cdot N1$ and $rU7\cdot N3$ CSs of hp-GU-24 (25 °C and pH 6.9). Errors in all RD-derived fitted parameters (for example, $\Delta\omega$) reflect experimental uncertainty (one s.d.) from the weighted global fit (see Methods). **b**, RD-derived CSs (referenced to GS WB) are shown for the ES of xpt-G riboswitch ($rU17\cdot N3$ and $rU69\cdot N3$) at 25 °C and pH 7.9. **c**, 2D [^{15}N , ^{13}C] HMQC spectra of rUTP showing CS of rUTP-N3 induced upon deprotonation. The spectra is rotated by 90°, to depict ^{15}N CS along x -axis for visualization purposes. Red circles on inset structure highlight measured resonances (C4 and N3). **d**, 2D [^{15}N , ^{13}C] HMQC spectra of rGTP showing CS of rGTP-N1 induced upon deprotonation. The spectra is rotated by 90°, to depict ^{15}N CS along x -axis for visualization purposes. Red circles on inset structure highlight measured resonances (C6 and N1). **e**, RD-derived hp-GU-20 $rG\cdot rU$ ES1 (blue) and ES2 (green) CSs are shown as a function of temperature and pH for both $rG16\cdot N1$ (square) and $rU5\cdot N3$ (circle).

(circle). **f**, Scheme used to calculate CSs using DFT. Idealized A-form RNA helix is generated to give a central $rG\cdot rU$ mispair (red) that is flanked by canonical $rG\cdot rC$ and $rA\cdot rU$ pairs, analogous to the hp-GU-24 construct. Residues are trimmed to 1-/9-methyl bases and $i + 1/i - 1$ pairs are frozen in place for subsequent geometry optimizations and CS calculations (see Methods). **g**, DFT-predicted CSs (referenced to an energy optimized WB geometry) are shown for various tautomeric and anionic configurations, where $rG^{enol}\cdot rU/rG\cdot rU^{enol}$ represents population weighted average CSs of $rG^{enol}\cdot rU$ (60%) and $rG\cdot rU^{enol}$ (40%). **h**, ^{15}N $rG\cdot N1$ and $rU\cdot N3$ CS comparison between RD-derived ES1 CSs and population weighted DFT-predicted CSs (60:40 vs. 80:20). **i**, Computational studies³² predict that the tautomeric pathway for a $rG\cdot rU$ pair can proceed via a planar $rG^+\cdot rU^-$ ion pair (charge delocalization is implied) that is highlighted by a network of five H-bonds. **j**, Pair geometry of an anionic $rG\cdot rU^-$ inverted wobble. Deprotonated $rU\cdot N3$ is highlighted in red (charge delocalization is implied).



Extended Data Figure 10 | dG•dT Misincorporation probabilities and correlation to WC-like excited states. **a**, Explicit dGTP•dT and dG•dTTP kinetic misincorporation and base substitution probabilities ($n = 53$) and associated errors^{46–48,72–77} (see Supplementary Discussion 8) are plotted against hp-GT dG•dT ES1 (blue squares) and ES2 (green triangles). The pK_a fit of ES2 probabilities to the Henderson–Hasselbalch equation (equation (4), see Methods) is shown as the green trend line. **b**, Red trend line shows the pK_a fit to

dGTP•dT misincorporation probabilities⁴⁷ from pH 6.5–8.6 to the Henderson–Hasselbalch equation. The fit was weighted using reported experimental errors and gave a reduced χ^2 of 3.56. **c**, **d**, Extrapolated dG•dT[–] ES2 probability (s.d. from the weighted global fit) is plotted against dGTP•dT (left) and dG•dTTP (right) misincorporation probabilities (errors as given)⁴⁷ from pH 6.5–9.5.

Crystal structure of the eukaryotic origin recognition complex

Franziska Bleichert¹, Michael R. Botchan² & James M. Berger¹

Initiation of cellular DNA replication is tightly controlled to sustain genomic integrity. In eukaryotes, the heterohexameric origin recognition complex (ORC) is essential for coordinating replication onset. Here we describe the crystal structure of *Drosophila* ORC at 3.5 Å resolution, showing that the 270 kilodalton initiator core complex comprises a two-layered notched ring in which a collar of winged-helix domains from the Orc1–5 subunits sits atop a layer of AAA+ (ATPases associated with a variety of cellular activities) folds. Although canonical inter-AAA+ domain interactions exist between four of the six ORC subunits, unanticipated features are also evident. These include highly interdigitated domain-swapping interactions between the winged-helix folds and AAA+ modules of neighbouring protomers, and a quasi-spiral arrangement of DNA binding elements that circumnavigate an approximately 20 Å wide channel in the centre of the complex. Comparative analyses indicate that ORC encircles DNA, using its winged-helix domain face to engage the mini-chromosome maintenance 2–7 (MCM2–7) complex during replicative helicase loading; however, an observed out-of-plane rotation of more than 90° for the Orc1 AAA+ domain disrupts interactions with catalytic amino acids in Orc4, narrowing and sealing off entry into the central channel. *Prima facie*, our data indicate that *Drosophila* ORC can switch between active and autoinhibited conformations, suggesting a novel means for cell cycle and/or developmental control of ORC functions.

The faithful replication of chromosomes relies on evolutionarily conserved initiator proteins to recruit ring-shaped helicases to DNA in a cell-cycle-regulated manner (reviewed in refs 1–3). Replication initiators belong to the AAA+ protein superfamily, a large group of multi-subunit nucleotide hydrolases that function as motors or molecular switches in many cellular processes⁴. AAA+ NTPases assemble into homo- or hetero-oligomeric complexes that actively alter the conformation or position of client macromolecules in response to ATP binding and hydrolysis.

In eukaryotes, replication initiation is promoted by the six-subunit ORC assembly^{5,6}. Five of the six subunits of ORC (Orc1–5) retain AAA+ modules^{4,7–9}, while the sixth (Orc6) is composed of tandem cyclin-box folds similar to transcription factor IIB (TFIIB)^{10,11}. During initiation, ORC binds replication origins, recruiting another AAA+ ATPase, Cdc6, to DNA in a nucleotide-dependent manner (reviewed in refs 3, 12). The DNA-bound ORC•Cdc6 complex in turn recruits the MCM2–7 replicative helicase and its associated Cdt1 chaperone to origins, promoting the loading of MCM2–7 complexes onto DNA (reviewed in refs 3, 13).

So far, many of the molecular mechanisms by which ORC assembles and operates have remained enigmatic. To better understand ORC function, we therefore determined the crystal structure of the *Drosophila* complex to 3.5 Å resolution. The structure highlights a domain-swapped organization for ORC and captures the complex in an unanticipated, autoinhibited conformation. Analysis of the structure leads to a revised model for DNA binding and proposed ORC•MCM2–7 contacts, and adds to a stepwise series of assembly and conformational intermediates that help account for how the complex acts during the early stages of replication initiation.

Crystal structure of the ORC hexamer

Sequence analyses had indicated that the Orc1–5 subunits would share a domain architecture similar to that of archaeal Orc proteins, with an AAA+-type ATPase fold fused to at least one carboxy (C)-terminal

winged-helix (WH) DNA-binding domain (Fig. 1a)^{7,8}. For its part, the Orc6 C terminus has been reported to bind to ORC1–5 through a domain insertion in Orc3, leaving its TFIIB-like domain conformationally independent of the ORC core (Fig. 1a)¹⁴. For crystallizing *Drosophila* ORC, we designed a ‘trimmed’ construct lacking the flexible amino (N)-terminal extensions of Orc1, Orc2 and Orc3 (ref. 14), and the Orc6 TFIIB region (Fig. 1a). Neither modification interfered with ORC assembly, nor did they affect the overall architecture of ORC (Extended Data Fig. 1a–c). This ORC core (referred to as ORC hereafter) crystallized in space group *I*222 with one Orc1–6 heterohexamer per asymmetric unit. The structure was phased by single-wavelength anomalous dispersion and refined to 3.5 Å with $R_{\text{work}}/R_{\text{free}}$ values of 0.22/0.26 (Extended Data Fig. 2a–c and Extended Data Table 1).

AAA+ and WH domains interlock within the ORC body

The ORC structure shows that the complex forms a lopsided, two-tiered ring with a cashew-shaped protuberance off of one edge (Fig. 1b, c and Supplementary Video 1). Orc1–5 comprise the ring body, which bears a prominent central channel, while a large domain insertion in Orc3 forms a bi-lobed, α -helical extension that engages a short α -helix formed at the Orc6 C terminus. In contrast to models based on prior electron microscopy (EM) reconstructions^{14,15}, the AAA+ subunits are arranged in the order of Orc1–Orc4–Orc5–Orc3–Orc2, thus revising the placement for Orc2 and Orc3 within the pentameric ORC ring (Fig. 1b).

In the structure, Orc1–5 each comprises one AAA+ fold, followed by a single C-terminal WH domain (Extended Data Fig. 3). The AAA+ and WH regions co-associate, but are segregated between the two ring tiers (Fig. 1b, c and Supplementary Video 1). Interestingly, the collar of WH domains is rotationally offset from the AAA+ domains, leading to a domain-swapped organization wherein, apart from Orc2, the WH domain of one subunit packs against the AAA+ domain of its adjoining partner (Fig. 1c). Domain swapping is facilitated by long linkages between the AAA+ and WH modules, a region known to be conformationally

¹Department of Biophysics and Biophysical Chemistry, Johns Hopkins School of Medicine, Baltimore, Maryland 21205, USA. ²Department of Molecular and Cell Biology, University of California Berkeley, Berkeley, California 94720, USA.

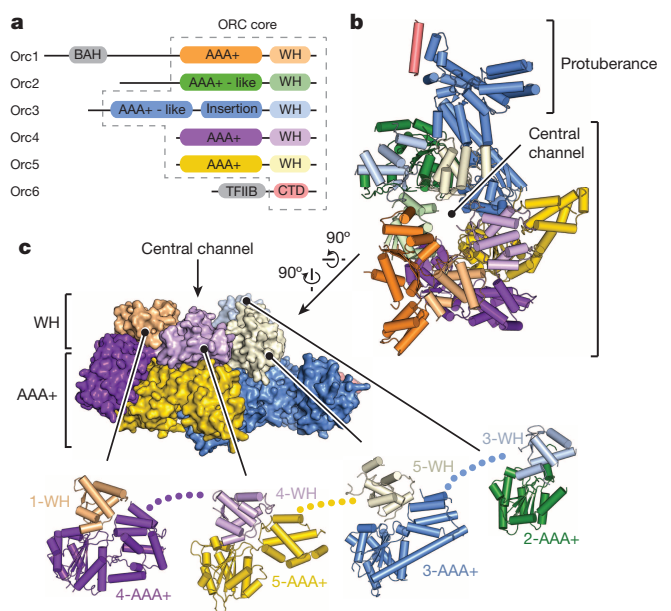


Figure 1 | Structure of *Drosophila* ORC. **a**, Domain organization of ORC subunits. Dashed lines demarcate the ORC core used for crystallization (TFIIB and CTD, transcription factor-II-like and C-terminal domains in Orc6; BAH, bromo-adjacent homology domain in Orc1). **b**, Crystal structure of ORC. Domains of individual subunits are coloured as in **a**. **c**, Side view of ORC (surface) highlighting the two-tiered, domain-swapped organization of the ORC body (Orc1 AAA+ and Orc2 WH domains, as well as N-terminal Orc2 residues (built as polyalanine) are not shown). Exploded view (cartoon) showing the packing of WH domains against adjacent subunits' AAA+ regions.

flexible in archaeal Orc homologues (Extended Data Fig. 3)^{16–19}. Other deviations from archetypal AAA+ structures include the absence of a small α -helical 'lid' subdomain in the AAA+ fold of Orc2, a larger-than-normal lid in Orc4 and a complete domain insertion in the lid of Orc3; the last augmentation forms the protuberance that binds Orc6 (Extended Data Figs 3 and 4 and Supplementary Discussion). Globally, the interdigitated interactions between the WH and AAA+ collars, as well as the extensive protein–protein contacts between adjoining WH domains, indicate that the WH domains are an important determinant of ORC stability (5,660 Å² of total surface area are buried between all WH domains, and 15,052 Å² between the WH and AAA+ tiers).

DNA binding elements in the WH domain collar

In archaeal Orc homologues, the WH element recognizes origin sequences by using a helix–turn–helix motif and a β -hairpin 'wing' to interact with the adjacent major and minor grooves of double-stranded DNA^{16,17,20,21} (Fig. 2a). Given the conservation between archaeal Orcs and eukaryotic Orc1–5 proteins, we anticipated that the WH domains of ORC would

bind DNA in a similar manner. However, the second α -helix of the helix–turn–helix in the WH domain (corresponding to the DNA 'recognition helix') is buried against the AAA+ tier in all subunits but Orc2 (Fig. 2b). This arrangement leaves the β -hairpin wings of Orc1, Orc4, Orc5 and Orc3 solvent exposed, which in turn co-localize to form a portion of the interior surface within the central channel in the ORC body (Fig. 2b, c). Given the extensive contacts between the WH and AAA+ tiers, sequestration of the recognition helices seems necessary to maintain ORC integrity. Thus, certain aspects of DNA recognition by the WH domains of ORC probably differ from the approach used by archaeal Orcs.

Quasi-canonical AAA+ domain packing within ORC

AAA+ ATPases tend to oligomerize into closed-ring, open-ring or helical assemblies^{4,9,22}. On the basis of phylogenetic, biochemical and EM data, ORC has been proposed to follow a similar trend^{7,8,14,23,24}. The ORC structure shows that, of the various AAA+ domain interactions, the one between Orc4 and Orc5 is most like that seen in typical AAA+ systems, whereby the two subunits associate to form a bipartite nucleotide-binding site at the inter-protomer interface (Fig. 3a and Extended Data Fig. 5a). Metazoan Orc4 manifests conserved Walker A and B ATPase motifs (GKT and D(D/E), respectively), which in related AAA+ ATPases typically contact both β - and γ -phosphates of bound nucleotide and a catalytically important Mg²⁺ ion. Interestingly, Orc5 donates a well-conserved 'arginine-finger' (Arg144) into the Orc4 active site (Extended Data Fig. 5a, b). Moreover, although Orc4 does not possess a typical 'Sensor II' arginine (an amino-acid motif in AAA+ ATPases that often aids hydrolysis), its Walker A region does retain a conserved basic amino acid (Arg58) (Extended Data Fig. 5c) that occupies an analogous position (Extended Data Fig. 5a). Overall, the observed structural organization and conservation of catalytically important residues in the Orc4•Orc5 interface raises the possibility that, in some organisms, Orc4 might be able to support some level of ATP turnover, conformationally modulate ORC function in response to ATP, or require nucleotide to promote ORC stability (as has been observed for the human complex^{25,26}). Any such activity probably varies between species, however, as *Saccharomyces cerevisiae* Orc4 neither contains a conserved Walker A motif nor has been found to bind/hydrolyse ATP²⁷.

Other AAA+/AAA+ interactions within ORC deviate to a greater or lesser extent from the canonical packing arrangements exemplified by Orc4•Orc5. For example, the Orc5•Orc3 interface is relatively open, with fewer contacts between AAA+ modules (Fig. 3a and Extended Data Fig. 5d). Although Orc5 possesses a canonical Walker A motif in most eukaryotes, the same is not true for its Walker B motif, nor is there a candidate Orc5 Sensor II residue or an Orc3 arginine finger. These structural features are consistent with biochemical observations showing that Orc5 binds ATP but lacks ATP hydrolysis activity^{26–30}. By comparison, residues for nucleotide binding and hydrolysis are not conserved at the Orc3•Orc2 AAA+/AAA+ interface; hence, the Orc2

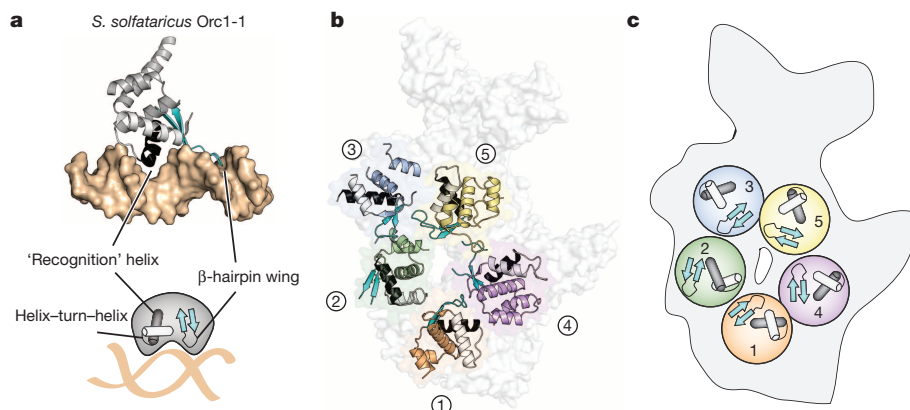


Figure 2 | Eukaryotic ORC and archaeal Orc WH domains. **a**, WH domain–DNA interactions in archaeal Orc1-1 (PDB accession number 2QBY chain A)¹⁶. The WH domain (grey) uses a helix–turn–helix (black and white) motif and a β -hairpin wing (cyan) motif to engage DNA (tan surface). **b**, Disposition of WH domains in ORC. The ORC WH domains (excepting Orc2) form a collar in which the recognition helices (black) are buried and the β -hairpin wings form an exposed portion of the central channel (the Orc1 AAA+ domain is not shown). The helix–turn–helix motifs and β -hairpin wings of ORC subunits are coloured as in **a**. **c**, Schematic of the helix–turn–helix and β -hairpin motifs in ORC.

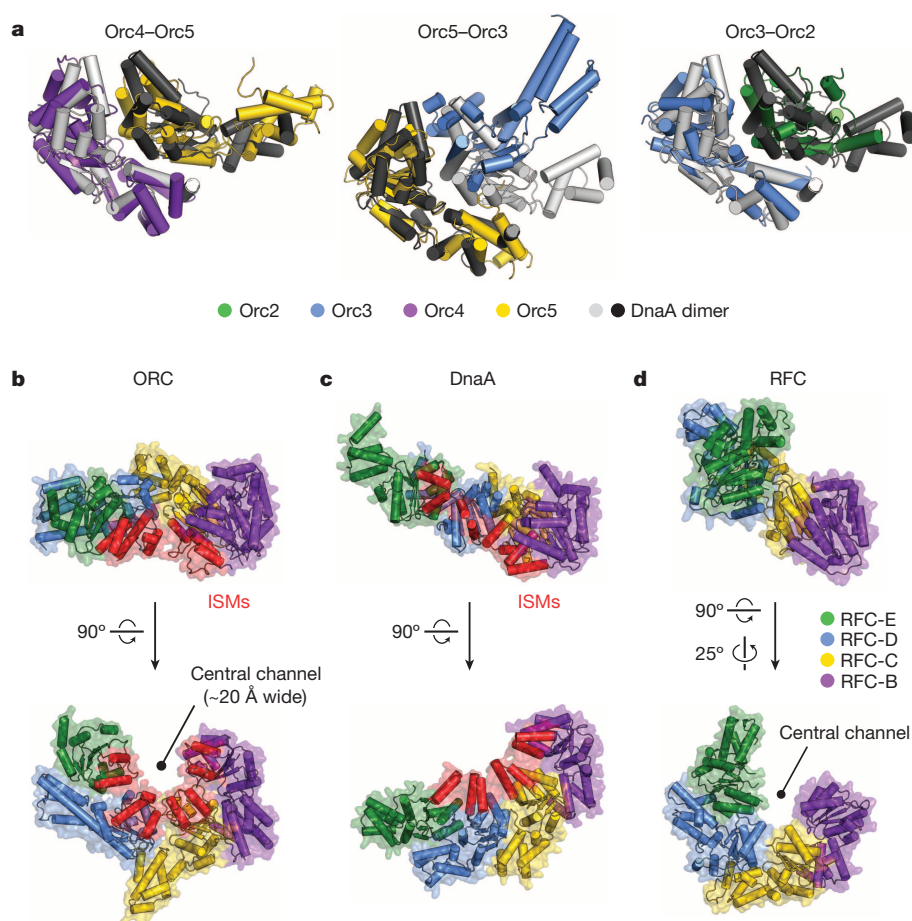


Figure 3 | AAA+/AAA+ domain interactions in the ORC. **a**, Pairwise AAA+ interactions in the Orc2–Orc5 oligomer compared with canonical AAA+ interactions in an ATP-bound DnaA dimer (PDB accession number 2HCB²²). Superpositions were performed using the AAA+ domain of the left-most protomer of a given ORC subunit pair and DnaA homodimer. **b–d**, Comparison of AAA+ domain assemblies (to scale) for Orc2–5 (**b**), ATP-assembled DnaA (PDB accession number 2HCB²²) (**c**) and replication factor C (RFC, PDB accession number 1SXJ³¹) (**d**). Upper panels depict side views, with the purple AAA+ domains of DnaA and RFC co-oriented as per the Orc4 AAA+ region. Lower panels show views down the central channel of ORC and RFC, or slightly offset from the helical axis of DnaA. The initiator specific motifs (ISMs, red) of Orc2–5 and DnaA are highlighted (RFC lacks this element). Only the AAA+ domains of ORC, DnaA and RFC are shown; the RFC-A subunit is omitted. Colouring for ORC is maintained as in **a**, with sequential subunits in DnaA and RFC coloured accordingly.

and Orc3 AAA+ domains and their interactions appear to play a predominantly structural part in ORC assembly.

Consistent with its blend of typical and atypical local AAA+/AAA+ interactions, the global organization between the ORC AAA+ domains is also mixed. Of AAA+ proteins, ORC is most closely related phylogenetically to other replication initiators, including bacterial DnaA, and to DNA polymerase clamp loaders⁹. In comparing the ORC ATPase assembly with these systems, it can be seen that the Orc4•Orc5 and Orc3•Orc2 AAA+ folds are organized similarly to protomers in ATP-assembled DnaA (Fig. 3a). By contrast, the AAA+/AAA+ organization at the Orc5•Orc3 interface is relatively open. Consequently, the clean helical symmetry observed in ATP-assembled DnaA oligomers²² is broken within the Orc5•Orc3 junction of ORC, a configuration that creates a cracked ring-like architecture reminiscent of clamp loaders³¹ (Fig. 3b–d). Despite the somewhat more planar arrangement of ATPase folds in ORC, the Orc2–5 initiator specific motifs (or ‘ISMs’, an α -helical modification that both distinguishes AAA+ family replication initiators and binds DNA directly^{8,16,17,22,32}) nonetheless cluster together, forming a shallow, quasi-spiral shaped set of ‘threads’ that line the interior of the ~20 Å wide central ORC channel (Fig. 3b). Thus, the structural features of ORC are a hybrid of both clamp loader and prokaryotic initiator systems.

An unanticipated Orc1 conformation

In addition to Orc4 and Orc5, one other ORC subunit also known to bind ATP is Orc1 (refs 27, 30). Indeed, Orc1 serves as the major source of ATPase activity in ORC and requires a conserved arginine residue from Orc4 for catalytic activity^{27–29,33}. Given this activity, we expected, as with Orc2–Orc5, to see relatively canonical AAA+ interactions between Orc1 and Orc4 in the structure. Surprisingly, the AAA+ domain of Orc1 is completely disengaged from Orc4, owing to a buckling of the

linker at a single region between the Orc1 ATPase and WH folds (residues 819–821) that gives rise to a ~105° out-of-plane rotation from the Orc2–5 AAA+ domains (Fig. 4a). Although this Orc1 AAA+/WH domain juxtaposition is unique compared with other subunits in the structure (Extended Data Fig. 3), it is similar to one of the conformations adopted by an ADP-bound archaeal Orc homologue, *Aeropyrum pernix* Orc2 (ref. 18) (Extended Data Fig. 6a, b). Notably, the movement of Orc1 does not considerably affect the relative arrangement of its two AAA+ subdomains, which is maintained as in Orc3–5 and archaeal Orcs (Extended Data Fig. 6c–e). The *en bloc* reorientation of Orc1 appears stabilized by contacts between the Orc1 AAA+ domain and the WH domains of Orc1, Orc2 and Orc3, together burying a total of 4,256 Å² at the interface (Fig. 4b and Extended Data Fig. 6f–h).

A consequence of the disposition of Orc1 within the complex is that its nucleotide binding cleft resides ~40 Å away from the arginine finger of Orc4. Importantly, comparison of the crystal structure with a prior three-dimensional EM reconstruction of ATP γ S-bound *Drosophila* ORC¹⁴ shows excellent agreement between the two models (Fig. 4c and Supplementary Video 2), indicating that the Orc1 conformation in the crystal corresponds to the predominant state of the complex in solution. Moreover, co-crystallization of ORC with the ATP analogue ATP γ S, while showing clear density for nucleotide binding to the Orc1, Orc4 and Orc5 AAA+ folds (Extended Data Fig. 7a–c), recapitulates the configuration seen in *apo*-ORC. Together, these data indicate that *Drosophila* Orc1 must undergo a large structural change to support ATPase activity, but that ATP binding is itself insufficient to drive such a rearrangement in most ORC particles.

Implications of the structure for DNA binding by ORC

In the ORC structure, the central channel within the body of the complex is formed by both the Orc2–5 ISMs and the β -hairpin wings of

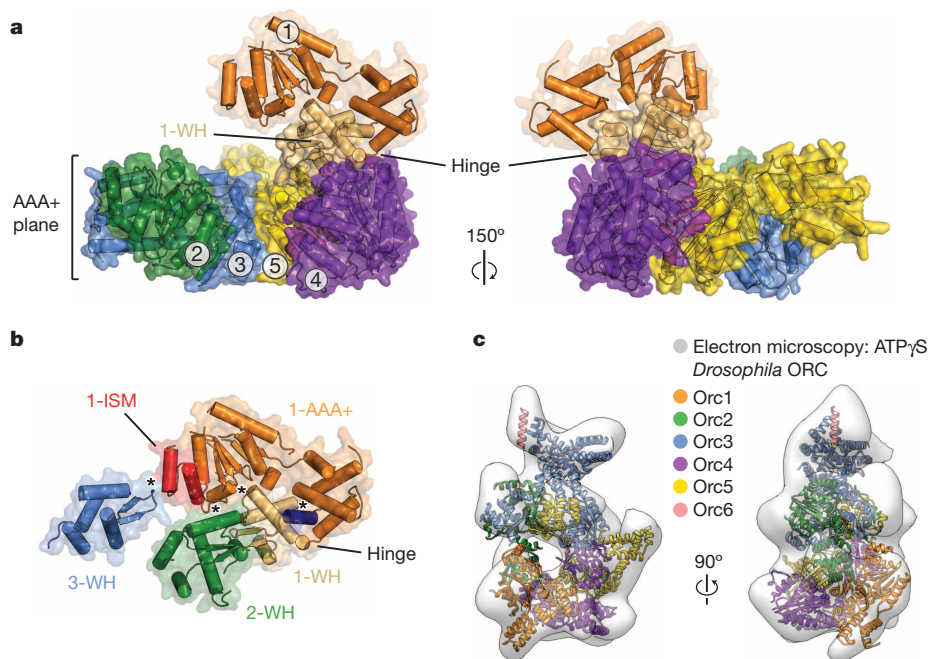


Figure 4 | An unanticipated but naturally occurring Orc1 conformation. **a**, The Orc1 AAA+ domain is disengaged from Orc4 and sits above the plane of the Orc2–Orc5 AAA+ oligomer. Only the AAA+ domains of Orc1–5 and the WH domain of Orc1 are depicted; the hinge point between the Orc1 AAA+ and WH regions is indicated. **b**, The Orc1 conformation is stabilized by contacts (asterisks) between its AAA+ domain and the WH domains of Orc1–3 (the C-terminal helix of Orc1 is in dark blue). See also Extended Data Fig. 6f–h. **c**, Docking study showing that the ORC crystal structure matches the three-dimensional EM volume of *Drosophila* ORC (EMD-2479 (ref. 14)). See also Extended Data Fig. 8c and Supplementary Video 2.

Orc1 and Orc3–5. In archaeal Orcs, these two elements both bind to duplex DNA^{16,17}. To investigate if the ORC central channel could accommodate any of the DNA interactions typified by archaeal Orcs, we superposed the DNA-bound crystal structure of *Sulfolobus solfataricus* Orc1-1 (ref. 16) onto Orc4 (after Orc1, Orc4 is most closely related to archaeal Orcs). Notably, superpositioning of the AAA+ domains of the two proteins (Fig. 5a) resulted in a placement for DNA that aligns the duplex coaxially with the central ORC channel (Fig. 5b). Inspection of the resultant ORC•DNA model not only reveals that the quasi-spiral formed by the Orc2–5 ISMs approximates that of the docked duplex, but that the β -hairpin wings of the Orc1 and Orc3–5 WH domains also reside in a position where they can access the nucleic-acid segment (Fig. 5b).

The superpositioning between DNA-bound archaeal Orc1-1 and the DNA-free *Drosophila* ORC imaged here has important implications for understanding how the eukaryotic initiator engages origin regions. One is that ORC probably binds DNA using a mechanism similar to that of sliding clamp loaders, which encircle primer-template junctions^{34,35}. This binding mode is congruent with a recent proposal based on EM analysis of an ORC•Cdc6•Cdt1•MCM2–7 complex, which posits that ORC helps align the ring of an MCM2–7 complex around DNA²⁴. The apparent sequence-specificity of DNA binding by *S. cerevisiae* ORC⁵ (in contrast to metazoan ORC^{36,37}) probably results from specific interactions between amino-acid side chains in the ORC channel and nucleotide bases in specific autonomously replicating sequence elements.

At present it is unclear from the structure why *Drosophila* ORC prefers to bind negatively supercoiled DNA over linear segments³⁶, a preference also reported for *Schizosaccharomyces pombe* ORC³⁸. Archaeal Orcs are known to underwind DNA upon binding^{16,17}, suggesting that the eukaryotic ORC AAA+ and WH domains may cooperate to do likewise; topological changes in DNA structure induced by *S. pombe* and human ORC are consistent with such an interpretation^{38,39}. Alternatively, if ORC were to bind B-form DNA, conformational transitions within the ORC body and its associated DNA-binding elements would be required to accommodate the DNA duplex. In this instance, specificity for negatively supercoiled substrates could arise from the relative positioning of the TFIIB-like DNA-binding elements in Orc6 (refs 10, 11, 40) and the DNA-recognition elements in the central ORC channel.

ORC activation and implications for MCM2–7 loading

Modelling indicates that the DNA binding elements in the central ORC channel can encircle a DNA duplex; however, certain features of the complex would appear to preclude *Drosophila* ORC from doing so in the state seen both here and by EM. For example, passage of the DNA through the entirety of the central channel is prevented by a constriction formed by both the Orc2 WH domain and the Orc1 AAA+ fold (Fig. 1b). Similarly, although the ATPase region of the ORC ring is cracked open, the Orc2 WH domain and the Orc1 AAA+ fold occlude

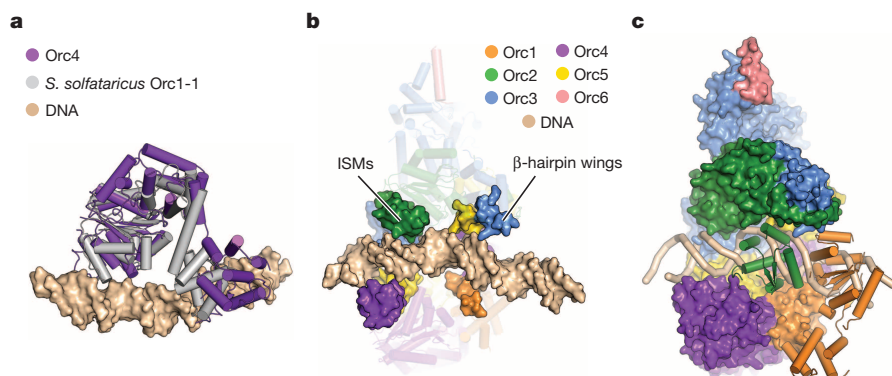


Figure 5 | The central channel in ORC probably binds DNA. **a**, Structural alignment of Orc4 onto DNA-bound *S. solfataricus* Orc1-1 (PDB accession number 2QBY chain A¹⁶). **b**, Reciprocal superposition of DNA-bound archaeal Orc1-1 onto Orc4 in the context of ORC aligns the DNA duplex with the central channel of ORC. The ISMs of Orc2–5 and β -hairpin wings of Orc1 and Orc3–5 (surface) form two semicircular, quasi-spiral collars around the docked DNA (Orc1 AAA+ and Orc2 WH domains are omitted; the remainder of ORC is shown as transparent cartoons). **c**, The DNA path (tan ribbon) through the central channel is partly blocked by the Orc1 AAA+ (orange cartoon) and the Orc2 WH (green cartoon) domains. The remainder of ORC is shown as surface representation. Colouring as in **b**.

this crack, thereby preventing the lateral entry of DNA into the central channel from the side of the complex (Figs 1b and 5c). This observation probably helps explain the weak effect that nucleotide has on DNA binding by *Drosophila* ORC purified from fly embryos or recombinant sources^{28,36}. It is interesting to note that by preventing DNA binding, the placement of the Orc1 ATPase region and the Orc2 WH fold also blocks the known interaction site for another critical component of replication initiation, Cdc6 (refs 23, 24). Overall, the simplest interpretation of the ORC conformation imaged here is that it corresponds to a naturally autoinhibited form of the complex, and that in some organisms, only a fraction of the total ORC pool that can be obtained from asynchronously dividing cells may be capable of productively altering its interactions with DNA in response to ATP.

If *Drosophila* ORC first assembles into an inactive form, then what manner of transition might push the complex into a new state in which its ATPase region is now competent to bind DNA (or Cdc6)? Insights into a simple structural rearrangement that could support such a switch can be gleaned from what is known about archetypal AAA+ ATPase organization. Using the Orc4•Orc5 interaction seen in the crystal structure as a template, we generated a model for the expected arrangement of a functional Orc1•Orc4 AAA+ ATPase centre by swivelling the Orc1 AAA+ fold around a single hinge point in the linker region just before its WH domain (Supplementary Video 1). The resultant model not only restores expected AAA+ interactions between the Orc1 active site and the Orc4 arginine finger (Fig. 6A and Extended Data Fig. 7d), but also both removes the Orc1-mediated blockage of the putative path for DNA in the central channel and co-aligns the Orc1 ISM with the ISM helix formed by Orc2–Orc5 (Fig. 6B and Supplementary Video 1). Docking of the rearranged model into the cryo-EM densities of *S. cerevisiae* ORC^{23,24} shows a reasonable fit for *Drosophila* ORC containing the repositioned Orc1 AAA+ domain, and further reveals that a region of EM density, which extends from the centre of the ORC body, actually corresponds to DNA (Extended Data Fig. 8a, b). In accord with a two-state model, the EM density for *S. cerevisiae* ORC cannot

accommodate the AAA+ domain of Orc1 in the state imaged crystallographically (Extended Data Fig. 8a), nor can the *Drosophila* ATPγS-ORC EM volume¹⁴ account for a remodelled Orc1 conformation for ORC (Extended Data Fig. 8c and Supplementary Video 2).

Collectively, the structure and analysis presented here helps to define a framework for understanding how *Drosophila* ORC interfaces with its partner proteins and DNA during the initial stages of MCM2–7 loading (Fig. 6C). In this scheme, ORC would start off in an ATP-bound but autoinhibited form—either dissociated from chromatin or in a chromatin-bound state via secondary binding sites/partners—that is restricted in its ability to either bind DNA within its central channel or bind Cdc6 to its ring. Conversion of this state into an activated configuration would involve the *en bloc* movement of the Orc1 ATPase domain (Supplementary Video 1), allowing Orc1 to engage the arginine finger of Orc4, and unlatching the Orc2 WH domain to open a gap in the Orc1–5 ring. Once open, DNA would bind to the ISM and β-hairpin elements in the central ORC channel, after which Cdc6 would dock into the Orc1/Orc2 gap, trapping DNA within the centre of the complex. One prediction of this model is that Cdc6 should bind to ORC using its ATPase centre to engage an arginine finger in Orc1 (*Drosophila* residue Arg734). After formation of a ternary ORC•DNA•Cdc6 complex, the WH domains of Cdc6 and Orc2 would be expected to engage the AAA+ folds of Orc1 and Cdc6, respectively, creating a circuit of WH domains (and their associated β-hairpin elements) to help lock the complex into place. Interestingly, these structural findings and analyses indicate that, rather than using its AAA+ domains to bind MCM2–7 (ref. 24), which requires an inverted order of ATPase site-arginine finger interactions around the ORC ring (Extended Data Fig. 9a, b), ORC probably uses its WH domain collar instead (Extended Data Fig. 9c).

A major question still remaining is what event might trigger ORC rearrangement, or why ORC should exist in an autoinhibited state. On the basis of our EM and co-crystallization data, ATP binding alone is incapable of efficiently driving such a transition for most of the particles

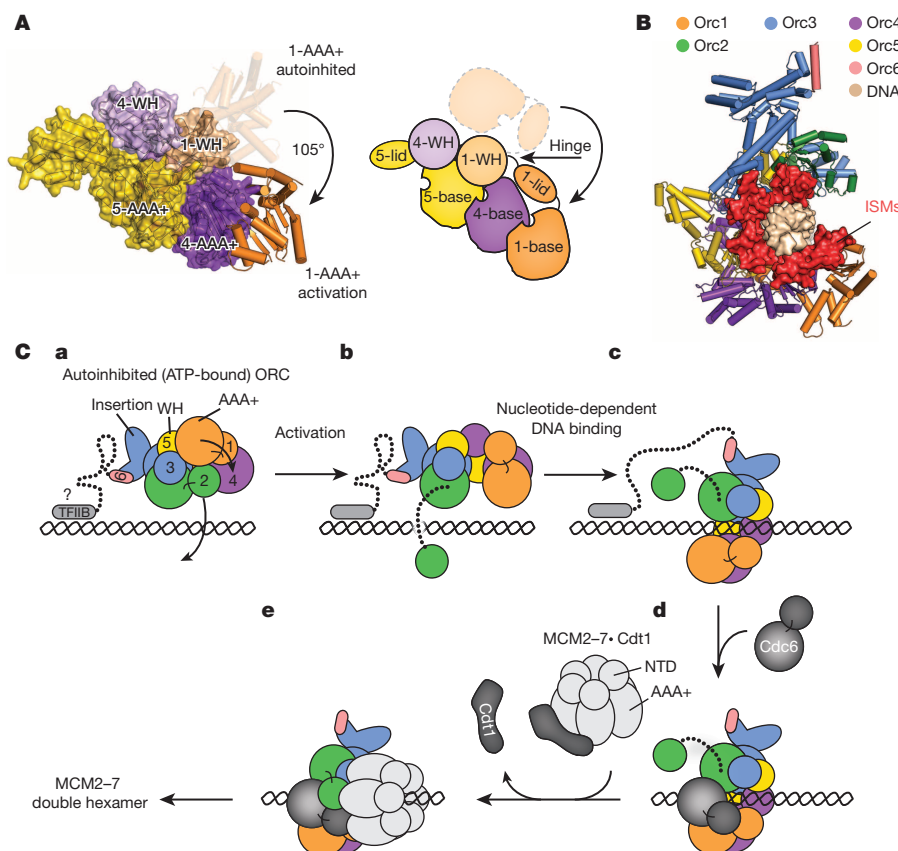


Figure 6 | Model for ORC activation and its functional consequences.

A, An approximate 105° rotation of the Orc1 ATPase fold about a hinge point within the AAA+/WH linker region juxtaposes the Orc1 active site with the arginine finger surface of Orc4 (only Orc1, Orc4 and the Orc5 AAA+ domain are depicted, both as a molecular model and as a schematic). **B**, Repositioning of Orc1 extends the circuit of ISMs around the prospective DNA binding site of the central channel. The Orc2 WH domain is omitted; DNA is modelled as per Fig. 5 and shown as surface. **C**, Model for ORC activation and subsequent Cdc6 and MCM2–7 recruitment/loading. **a**, Prospective recruitment of autoinhibited *Drosophila* ORC to DNA by elements such as the TFIIB domain of Orc6. **b**, Activation of ORC promotes productive Orc1•Orc4 interactions and Orc2 WH domain repositioning, exposing a gap in the ORC ring. **c**, DNA binds to the central channel of ORC through ISM/β-hairpin wing interactions. **d**, Cdc6 binds between Orc2 and Orc1, trapping DNA inside the complex. **e**, ORC•Cdc6 binds to and loads one MCM2–7 hexamer (aided by Cdt1) using the WH domain face of ORC•Cdc6. During loading, two MCM2–7 particles are assembled into a double hexamer^{47,48}.

in a purified ORC preparation. Interestingly, phosphatase treatment of *Drosophila* ORC stimulates DNA binding⁴¹, suggesting that removal of one or more post-translational marks might help convert ORC to an active form. Moreover, metazoan ORC associates with chromatin in a cell-cycle-dependent manner (reviewed in refs 1, 42), and the targeting of ORC to chromosomes in metazoans (and in fission yeast) is known to require protein–DNA contacts distinct from those in budding yeast, such as the TFIIB-homology domain in metazoan Orc6 (refs 10, 40) or the AT-hook in *S. pombe* Orc4 (ref. 43). The action of these elements, together with the recognition of nucleosomes by the N-terminal BAH domain of Orc1 (refs 44–46), suggest that the formation of nucleotide-dependent contacts between the ATPase region of ORC and DNA may take place after the formation of initial ORC·chromatin encounter complexes. Moreover, metazoans need to stockpile ORC in oocytes, yet keep this pool of ORC from prematurely initiating DNA replication before fertilization; hence, ORC may be maternally stored in an inactive form during oogenesis. In these contexts, autoinhibition could provide a novel mechanism for regulating ORC's productive association with DNA in a cell-cycle-dependent and/or developmental stage-specific manner.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 26 October 2014; accepted 16 January 2015.

Published online 11 March 2015.

- Siddiqui, K., On, K. F. & Diffley, J. F. Regulating DNA replication in eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, a012930 (2013).
- O'Donnell, M., Langston, L. & Stillman, B. Principles and concepts of DNA replication in bacteria, archaea, and eukarya. *Cold Spring Harb. Perspect. Biol.* **5**, a010108 (2013).
- Bell, S. P. & Kaguni, J. M. Helicase loading at chromosomal origins of replication. *Cold Spring Harb. Perspect. Biol.* **5**, a010124 (2013).
- Neuwald, A. F., Aravind, L., Spouge, J. L. & Koonin, E. V. AAA+: a class of chaperone-like ATPases associated with the assembly, operation, and disassembly of protein complexes. *Genome Res.* **9**, 27–43 (1999).
- Bell, S. P. & Stillman, B. ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128–134 (1992).
- Diffley, J. F. & Cocker, J. H. Protein-DNA interactions at a yeast replication origin. *Nature* **357**, 169–172 (1992).
- Speck, C., Chen, Z., Li, H. & Stillman, B. ATPase-dependent cooperative binding of ORC and Cdc6 to origin DNA. *Nature Struct. Mol. Biol.* **12**, 965–971 (2005).
- Clarey, M. G. *et al.* Nucleotide-dependent conformational changes in the DnaA-like core of the origin recognition complex. *Nature Struct. Mol. Biol.* **13**, 684–690 (2006).
- Iyer, L. M., Leippe, D. D., Koonin, E. V. & Aravind, L. Evolutionary history and higher order classification of AAA+ ATPases. *J. Struct. Biol.* **146**, 11–31 (2004).
- Chesnokov, I. N., Chesnokova, O. N. & Botchan, M. A cytokinetic function of *Drosophila* ORC6 protein resides in a domain distinct from its replication activity. *Proc. Natl Acad. Sci. USA* **100**, 9150–9155 (2003).
- Liu, S. *et al.* Structural analysis of human Orc6 protein reveals a homology with transcription factor TFIIB. *Proc. Natl Acad. Sci. USA* **108**, 7373–7378 (2011).
- Costa, A., Hood, I. V. & Berger, J. M. Mechanisms for initiating cellular DNA replication. *Annu. Rev. Biochem.* **82**, 25–54 (2013).
- Yardimci, H. & Walter, J. C. Prereplication-complex formation: a molecular double take? *Nature Struct. Mol. Biol.* **21**, 20–25 (2014).
- Bleichert, F. *et al.* A Meier-Gorlin syndrome mutation in a conserved C-terminal helix of Orc6 impedes origin recognition complex formation. *eLife* **2**, e00882 (2013).
- Chen, Z. *et al.* The architecture of the DNA replication origin recognition complex in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **105**, 10326–10331 (2008).
- Dueber, E. L., Corn, J. E., Bell, S. D. & Berger, J. M. Replication origin recognition and deformation by a heterodimeric archaeal Orc1 complex. *Science* **317**, 1210–1213 (2007).
- Gaudier, M., Schuwirth, B. S., Westcott, S. L. & Wigley, D. B. Structural basis of DNA replication origin recognition by an ORC protein. *Science* **317**, 1213–1216 (2007).
- Singleton, M. R. *et al.* Conformational changes induced by nucleotide binding in Cdc6/ORC from *Aeropyrum pernix*. *J. Mol. Biol.* **343**, 547–557 (2004).
- Liu, J. *et al.* Structure and function of Cdc6/Cdc18: implications for origin recognition and checkpoint control. *Mol. Cell* **6**, 637–648 (2000).
- Dueber, E. C. *et al.* Molecular determinants of origin discrimination by Orc1 initiators in archaea. *Nucleic Acids Res.* **39**, 3621–3631 (2011).
- Robinson, N. P. *et al.* Identification of two origins of replication in the single chromosome of the archaeon *Sulfolobus solfataricus*. *Cell* **116**, 25–38 (2004).
- Erzberger, J. P., Mott, M. L. & Berger, J. M. Structural basis for ATP-dependent DnaA assembly and replication-origin remodeling. *Nature Struct. Mol. Biol.* **13**, 676–683 (2006).
- Sun, J. *et al.* Cdc6-induced conformational changes in ORC bound to origin DNA revealed by cryo-electron microscopy. *Structure* **20**, 534–544 (2012).
- Sun, J. *et al.* Cryo-EM structure of a helicase loading intermediate containing ORC–Cdc6–Cdt1–MCM2–7 bound to DNA. *Nature Struct. Mol. Biol.* **20**, 944–951 (2013).
- Ranjan, A. & Gossen, M. A structural role for ATP in the formation and stability of the human origin recognition complex. *Proc. Natl Acad. Sci. USA* **103**, 4864–4869 (2006).
- Siddiqui, K. & Stillman, B. ATP-dependent assembly of the human origin recognition complex. *J. Biol. Chem.* **282**, 32370–32383 (2007).
- Klemm, R. D., Austin, R. J. & Bell, S. P. Coordinate binding of ATP and origin DNA regulates the ATPase activity of the origin recognition complex. *Cell* **88**, 493–502 (1997).
- Chesnokov, I., Remus, D. & Botchan, M. Functional analysis of mutant and wild-type *Drosophila* origin recognition complex. *Proc. Natl Acad. Sci. USA* **98**, 11997–12002 (2001).
- Giordano-Coltart, J., Ying, C. Y., Gautier, J. & Hurwitz, J. Studies of the properties of human origin recognition complex and its Walker A motif mutants. *Proc. Natl Acad. Sci. USA* **102**, 69–74 (2005).
- Kong, D., Coleman, T. R. & DePamphilis, M. L. *Xenopus* origin recognition complex (ORC) initiates DNA replication preferentially at sequences targeted by *Schizosaccharomyces pombe* ORC. *EMBO J.* **22**, 3441–3450 (2003).
- Bowman, G. D., O'Donnell, M. & Kuriyan, J. Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* **429**, 724–730 (2004).
- Duderstadt, K. E., Chuang, K. & Berger, J. M. DNA stretching by bacterial initiators promotes replication origin opening. *Nature* **478**, 209–213 (2011).
- Bowers, J. L., Randell, J. C., Chen, S. & Bell, S. P. ATP hydrolysis by ORC catalyzes reiterative MCM2–7 assembly at a defined origin of replication. *Mol. Cell* **16**, 967–978 (2004).
- Kelch, B. A., Makino, D. L., O'Donnell, M. & Kuriyan, J. How a DNA polymerase clamp loader opens a sliding clamp. *Science* **334**, 1675–1680 (2011).
- Simonetta, K. R. *et al.* The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* **137**, 659–671 (2009).
- Remus, D., Beall, E. L. & Botchan, M. R. DNA topology, not DNA sequence, is a critical determinant for *Drosophila* ORC-DNA binding. *EMBO J.* **23**, 897–907 (2004).
- Vashee, S. *et al.* Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev.* **17**, 1894–1908 (2003).
- Houchens, C. R. *et al.* Multiple mechanisms contribute to *Schizosaccharomyces pombe* origin recognition complex-DNA interactions. *J. Biol. Chem.* **283**, 30216–30224 (2008).
- Gaczynska, M. *et al.* Atomic force microscopic analysis of the binding of the *Schizosaccharomyces pombe* origin recognition complex and the spOrc4 protein with origin DNA. *Proc. Natl Acad. Sci. USA* **101**, 17952–17957 (2004).
- Balasov, M., Huijbregts, R. P. & Chesnokov, I. Role of the Orc6 protein in origin recognition complex-dependent DNA binding and replication in *Drosophila melanogaster*. *Mol. Cell. Biol.* **27**, 3143–3153 (2007).
- Remus, D., Blanchette, M., Rio, D. C. & Botchan, M. R. CDK phosphorylation inhibits the DNA-binding and ATP-hydrolysis activities of the *Drosophila* origin recognition complex. *J. Biol. Chem.* **280**, 39740–39751 (2005).
- DePamphilis, M. L. Cell cycle dependent regulation of the origin recognition complex. *Cell Cycle* **4**, 70–79 (2005).
- Chuang, R. Y. & Kelly, T. J. The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proc. Natl Acad. Sci. USA* **96**, 2656–2661 (1999).
- Kuo, A. J. *et al.* The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier–Gorlin syndrome. *Nature* **484**, 115–119 (2012).
- Muller, P. *et al.* The conserved bromo-adjacent homology domain of yeast Orc1 functions in the selection of DNA replication origins within chromatin. *Genes Dev.* **24**, 1418–1433 (2010).
- Onishi, M. *et al.* Role of the conserved Sir3-BAH domain in nucleosome binding and silent chromatin assembly. *Mol. Cell* **28**, 1015–1028 (2007).
- Evrin, C. *et al.* A double-hexameric MCM2–7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl Acad. Sci. USA* **106**, 20240–20245 (2009).
- Remus, D. *et al.* Concerted loading of MCM2–7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the beamline scientists at beamlines 8.3.1 of the Advanced Light Source (Berkeley Lawrence National Laboratory), 23-ID-B of the Advanced Photon Source (Argonne National Laboratory) and X25 of the National Synchrotron Light Source (Brookhaven National Laboratory) for technical support with X-ray diffraction data collection, and the Nogales laboratory (University of California, Berkeley) for EM support. We also thank A. Fisher and Y. Li for assistance with insect cell cultures. This work was supported by the National Institutes of Health (GM071747 to J.M.B. and CA R37-30490 to M.R.B.) and by a fellowship from the UC Berkeley Miller Institute for Basic Research in Science (to F.B.).

Author Contributions F.B. performed all biochemical and crystallization experiments, collected X-ray diffraction data and determined the structure with guidance from J.M.B. All authors interpreted and discussed results, and wrote the manuscript.

Author Information Atomic coordinates and structure factors have been deposited in the RSCB Protein Data Bank under accession code 4XGC. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.M.B. (jberge29@hmi.edu) or M.R.B. (mbotchan@berkeley.edu).

METHODS

No statistical methods were used to predetermine sample size.

ORC construct design for crystallization. Initial crystallization trials with full-length *Drosophila* ORC did not yield crystals. To obtain a crystallizable complex, we therefore identified regions in ORC subunits likely to be unconserved, disordered or flexibly tethered to ORC by multiple sequence alignments and EM¹⁴. On the basis of these analyses, we designed a *Drosophila* ORC construct lacking the N-terminal regions that precede the AAA+ domains of Orc1 (amino acids 1–532), Orc2 (amino acids 1–265) and Orc3 (amino acids 1–46). These truncated subunits were found to assemble into a stable hexameric complex and yielded two-dimensional class averages very similar to full-length ORC when analysed by negative-stain EM, indicating that removal of the termini does not affect overall ORC architecture (Extended Data Fig. 1). In addition, we previously found that the TFIIB-like domain of Orc6 is flexibly tethered to ORC via the Orc6 C-terminal domain¹⁴; this region was also removed for ORC crystallization. All biochemical experiments described in the present paper were performed with this 'trimmed' ORC construct.

Cloning and baculovirus generation. ORC subunits Orc1 to Orc5 (Orc1: amino-acid residues 533–924; Orc2: amino-acid residues 266–618; Orc3: amino-acid residues 47–721; Orc4: amino-acid residues 1–459; Orc5: amino-acid residues 1–460) were cloned into a pFastBac-derived polycistronic BioBricks vector (University of California Berkeley MacroLab). A hexa-histidine (6× His) tag was added to the N terminus of Orc1 and a maltose binding protein (MBP) tag to the N terminus of Orc4, both followed by a tobacco etch virus (TEV) protease cleavage site. The C terminus of Orc6 (amino-acid residues 187–257) was cloned into a separate pFastBac vector. For Orc6 binding and crosslinking experiments, Orc6 (amino-acid residues 187–257) was cloned with an N-terminal 6× His tag into a ligation-independent-cloning (LIC)-compatible pFastBac vector¹⁴. Point mutations (Y225S, A236E and A236C) were introduced by site-directed mutagenesis and verified by DNA sequencing.

Bacmids were generated in DH10Bac cells and isolated as per the Bac-to-Bac protocol (Invitrogen Life Technologies). Sf9 cells were transfected with bacmid DNA using Cellfectin II (Invitrogen Life Technologies), also according to the manufacturer's instructions. Baculoviruses were amplified twice in Sf9 cells before infecting large-scale cultures for protein expression.

Expression and purification of ORC for crystallization. For expression, 8 l of High5 cells grown in spinner flasks were co-infected with two baculoviruses: the multibac virus expressing Orc1–5 and the baculovirus expressing Orc6. Forty-eight hours after infection, cells were harvested by centrifugation and ORC was purified as described in ref. 14 with a few modifications. Briefly, harvested cells were resuspended in 200 ml of lysis buffer (50 mM Tris-HCl (pH 7.8), 300 mM KCl, 50 mM imidazole (pH 7.8), 10% glycerol, 200 μM PMSF, 1 μg ml^{−1} leupeptin) and lysed by sonication. The lysate was clarified by centrifugation, treated with a 20% (NH₄)₂SO₄ precipitation on ice for 30 min, and re-centrifuged. ORC was purified by passing the supernatant solution over a 5 ml HisTrap HP nickel-affinity chromatography column (GE Healthcare), followed by amylose-affinity chromatography using a 20 ml column (New England Biolabs). Purification tags were removed from ORC by incubation with 6× His-tagged TEV protease⁴⁹ overnight at 4 °C. TEV protease and any uncleaved His-tagged Orc1-containing material was then removed by an additional nickel-affinity chromatography step in 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 50 mM imidazole (pH 7.8) and 10% glycerol. The flow-through was concentrated in 30K Amicon Ultra-15 concentrators (Millipore) and purified by gel filtration chromatography using a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) equilibrated in 50 mM Tris-HCl (pH 7.8), 300 mM KCl and 5% glycerol. Peak fractions were pooled and concentrated using 30K Amicon Ultra-15 concentrators (Millipore). Protein was used immediately thereafter for crystallography.

Crystallization. Before crystallization, ORC was dialysed overnight into crystallization buffer (50 mM Tris-HCl (pH 7.8), 200 mM KCl, 5% glycerol and 0.5 mM TCEP). Crystals grew by sitting-drop vapour diffusion as showers of small plates within a few hours upon combining equal volumes of 10 mg ml^{−1} ORC with reservoir solution (50 mM PIPES (pH 7.5), 100 mM ammonium acetate, 10–35 mM MgCl₂ and 2.5–5% PEG 20,000) at 19–22 °C. These crystals were used as a source for streak seeding to obtain larger crystals, which grew to their maximum size within 2 days. For cryo-protection, the mother liquor was stepwise exchanged (typically two steps for 3 h and 15 min, respectively) into 50 mM PIPES (pH 7.5), 25 mM KCl, 10 mM MgCl₂, 50 mM ammonium acetate, 4% PEG 20,000, 10% glycerol, 10% ethylene glycol, 20% xylitol and 0.25 mM TCEP. Crystals were then harvested by looping and flash-frozen in liquid nitrogen. (NB: during optimization, a subset of crystals was transferred to 4 °C and, after several days of equilibration, cryo-protected and harvested at this temperature; this procedure seemed to slightly increase the number of usable crystals.) For phasing, crystals were soaked into cryo-protecting solution containing 1 mM GdCl₃ for 3–4 h before harvesting.

Additional heavy-atom-derivatized crystals were obtained by incubating crystals in cryo-protecting solution containing 100 μM ethyl mercuric phosphate for 3 h; data from these soaks were used to identify metal-binding sites for confirming amino-acid registers but were not used for phasing.

Crystallization and crystal harvesting for ATPγS-bound ORC was performed as described for apo-ORC but with the following modifications: (1) ORC was dialysed into crystallization buffer containing MgCl₂ (50 mM Tris-HCl (pH 7.8), 200 mM KCl, 5% glycerol, 0.5 mM TCEP, 5 mM MgCl₂); (2) before crystallization, ATPγS was added to ORC to a final concentration of 1 mM; and (3) for harvesting, 0.5 mM ATPγS was added to the cryo-protectant to prevent dissociation of the nucleotide.

Data collection and structure determination. The diffraction quality of individual crystals varied greatly, necessitating the screening of many hundreds of different crystals to identify acceptable targets for data collection. Crystal screening was performed at beamlines 8.3.1 at the Advanced Light Source at Lawrence Berkeley National Laboratory, X25 at the National Synchrotron Light Source at Brookhaven National Laboratory and 23-ID-B at the Advanced Photon Source at Argonne National Laboratory. Native diffraction data sets ($\lambda = 1.0332$ Å) as well as single-wavelength anomalous dispersion (SAD) data sets for gadolinium ($\lambda = 1.71083$ Å) and sulphur ($\lambda = 1.7712$ Å, see below) were collected at the Advanced Photon Source 23-ID-B equipped with a MAR CCD detector. Although complete data sets were typically obtained by exposing multiple regions within a single crystal using the 'vector data collection' option to minimize radiation damage, the best native data set was collected at a single site from a crystal harvested at 4 °C. Data sets of ethyl mercuric phosphate-derivatized crystals were collected at the National Synchrotron Light Source X25 ($\lambda = 1.006$ Å) on a Pilatus 6M detector.

Diffraction data were processed with XDS^{50,51}. Merging with the program AIMLESS of the CCP4 software package⁵² (Extended Data Table 1) indicated that the crystals belonged to space group *I*222, with unit cell dimension of $a = 145.5$ Å, $b = 259.0$ Å and $c = 257.0$ Å for the best native crystal. The Gd-SAD data set was obtained by merging data from four different crystals. Despite slight non-isomorphism between crystals, merging data from multiple crystals substantially improved the anomalous signal, phases and electron-density interpretability compared with data sets collected from single crystals.

For initial phasing by SAD, gadolinium sites were identified with SHELXD⁵³. The strongest sites were then used as input into PHASER⁵⁴ as implemented in PHENIX⁵⁵ to find additional sites and to obtain initial phases to ~4 Å. Maximum likelihood density modification with RESOLVE⁵⁶ was used to break the phase ambiguity and to improve electron-density maps. At this point, experimental phases were next applied (again using PHENIX) to the native data set, which was of better quality than the gadolinium derivative. The resulting electron-density maps (at 4 Å resolution) allowed identification of all five AAA+ domains and four of the five WH domains, and revealed density for bulky side chains as expected for this resolution. In parallel to data processing with SHELXD/PHASER/RESOLVE, experimental phases and density maps were also calculated with SHARP and improved by solvent flipping in SOLOMON^{57,58}. Although slightly less featured, the SOLOMON electron-density maps were overall very similar to those obtained from RESOLVE but additionally revealed clear protein density in some regions that were poorly defined in RESOLVE density maps. Thus, while model building was performed predominantly into RESOLVE density maps, SOLOMON density maps were used as an additional guide to trace the main chain of the model. Phases were gradually improved by iterative cycles of model building, density modification with phase combination of experimental and model phases in RESOLVE, and phase extension to 3.7 Å.

Model building was initiated by automated searches using MOLREP^{52,59} and by manual docking using UCSF Chimera^{60,61} to place the AAA+ and WH domains of archaeal Orc1/Cdc6 (PDB accession numbers: 2QBY¹⁶; 1FNN¹⁹) and/or homology models for *Drosophila* ORC subunits (as generated by Phyre2 (ref. 62)) into electron-density maps. These docked structures were valuable reference points and facilitated tracing of most of the main chain; insights into likely subunit positions (from low-resolution EM studies of *Drosophila* ORC¹⁴), together with knowledge of the domain architecture of ORC subunits (from sequence predictions), allowed assignment of specific subunits to map density regions. Using COOT⁶³, a nearly complete model of ORC was manually built *de novo* into phase-combined and B-factor-sharpened RESOLVE density maps, guided by the topology of archaeal Orc1/Cdc6 AAA+ and WH domains, as well as by secondary structure prediction and multiple sequence alignments. The initial model was improved by iterative rounds of refinement in PHENIX (real-space, individual xyz, individual atomic displacement parameters), using secondary structure and (in early stages of building) experimental phase restraints, as well as stereochemistry and atomic displacement parameter weight optimization; subsequent rounds of model rebuilding were performed using COOT. During the course of refinement, a slightly higher-resolution and more complete native data set (to 3.5 Å) was obtained and used for

the final rounds of refinement, which also included refinement of TLS parameters. The final model contains the AAA+ and WH domains of Orc1 to Orc5, the Orc3 domain insertion and the conserved C-terminal helix of Orc6; an N-terminal region of Orc2 (preceding the AAA+ domain) was built as a poly-alanine model, since the amino-acid register for this region could not be assigned unambiguously. The final model was validated with MOLPROBITY⁶⁴ and has excellent geometry (MolProbity score 1.88), with no Ramachandran outliers and only a small fraction (1.9%) of rotamer outliers (Extended Data Table 1).

During the course of model building, several approaches were used to validate the sequence register of the various ORC chains. These included (1) using Hg-binding sites in ethyl mercuric phosphate-derivatized crystals as fiducials for cysteines and histidines, (2) using sulphur sites in native S-SAD data sets to verify the location of a subset of cysteines and methionines (the weak signal present in these data precluded the use of this information for phasing) and (3) conducting Orc3–Orc6 crosslinking experiments to confirm the register of the Orc6 C-terminal helix (Extended Data Fig. 4h). Hg and S sites were identified from log-likelihood-gradient maps calculated using the MR-SAD option in PHASER⁵⁴. Additionally, MR-SAD for sulphur sites also revealed the position of two ions with anomalous scattering properties at the wavelength of data collection ($\lambda = 1.7712 \text{ \AA}$). These ions showed clear density in experimental and $2F_o - F_c$ maps and were interpreted as a chloride ion in the P-loop of Orc5 and as a potassium ion bound to Orc2.

Once a satisfactory apo-ORC model was obtained, it was used as a search model for molecular replacement (using PHENIX-PHASER) to phase data collected from ORC co-crystallized with ATP γ S. The resulting solution (Z score = 86.7, log-likelihood-gain = 7747) revealed clear (more than $2-4\sigma$) positive difference density in the nucleotide binding clefts of Orc1, Orc4 and Orc5 that could accommodate ATP γ S (Extended Data Fig. 7a–c). Since the resolution of these crystals was limited to 4 Å, and since only small structural changes were observed throughout the remainder of ORC, we refrained from building and refining a model against this data set.

Structure analysis. Structural superpositions and docking into EM maps were performed using UCSF Chimera^{60,61}. Buried surface area at domain/subunit interfaces was calculated with PyMOL (PyMOL Molecular Graphics System, version 1.7.0.0, Schrödinger). Multiple protein sequence alignments were performed with MAFFT^{65,66} and conservation scores were calculated and mapped onto the structure with ConSurf⁶⁷. Sequence logos were generated with WEBLOGO⁶⁸. Figures were rendered using both PyMOL (The PyMOL Molecular Graphics System, version 1.7.0.0, Schrödinger) and UCSF Chimera^{60,61}.

Expression and purification of ORC1–5 and Orc6 for binding and crosslinking studies. ORC containing subunits ORC1–5 was expressed in High5 cells using a single virus with the multibac approach. Expression and purification were performed as described above for expression of ORC1–6, except that (1) 1 mM β -mercaptoethanol or 1 mM DTT was included in buffers during nickel- and amylose-affinity steps or gel filtration chromatography, respectively, and (2) 10% glycerol was maintained in all buffers throughout purification. For crosslinking experiments, 6 \times His and MBP tags were removed from ORC1–5, whereas for fluorescence anisotropy and pull-down experiments the affinity tags were left on ORC1–5 by omitting the TEV protease cleavage and subsequent nickel-affinity steps.

The CTD of Orc6 (residues 187–257) was purified as described previously¹⁴. Briefly, 6 \times His-tagged Orc6-CTD^{WT}, Orc6-CTD^{A236E}, Orc6-CTD^{A236C} and Orc6-CTD^{Y225S} constructs were expressed in High5 cells and purified by nickel-affinity chromatography in lysis buffer (50 mM Tris-HCl (pH 7.8), 600 mM KCl, 10% glycerol, 50 mM imidazole, 1 mM β -mercaptoethanol, 200 μ M PMSF and 1 μ g ml^{−1} leupeptin). During the second of two wash steps, the salt concentration was decreased to 300 mM KCl. Protein was eluted with 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 10% glycerol, 250 mM imidazole and 1 mM β -mercaptoethanol, concentrated in 3K Amicon Ultra-15 concentrators (Millipore), and further purified on a HiPrep 16/60 Sephacryl S-200 HR column (GE Healthcare) equilibrated in 50 mM Tris-HCl (pH 7.8), 300 mM KCl, 10% glycerol and 1 mM DTT. Peak fractions were pooled, concentrated in 3K Amicon Ultra-15 concentrators (Millipore) and stored at -80°C .

Fluorescence anisotropy. Orc6 proteins were N-terminally labelled with Alexa Fluor 488 5-SDP ester (Invitrogen Life Technologies) as described previously¹⁴. Binding reactions were performed for 30 min in a total volume of 140 μ l containing 30 nM fluorescently labelled wild-type Orc6-CTD, Orc6-CTD^{Y225S} or Orc6-CTD^{A236E}, and different concentrations of ORC1–5 (ranging from 122 pM to 1 μ M in twofold serial dilutions) in a buffer consisting of 50 mM Tris (pH 7.8), 300 mM KCl, 5% glycerol, 1 mM DTT and 0.1 mg ml^{−1} BSA. Forty microlitres of each reaction were transferred to 384-well plates in triplicates and fluorescence polarization was measured in a POLARstar Omega plate reader (BMG Labtech) with a 485 nm excitation filter and a 520 nm emission filter. For data analysis, fluorescence polarization measurements were converted into fluorescence anisotropy units, which ranged from 0.01 to 0.16. The mean fluorescence anisotropy

values obtained from three (for Orc6-CTD^{Y225S} and Orc6-CTD^{A236E}) or six (for wild-type Orc6-CTD) independent experiments were fitted to the quadratic binding equation:

$$A = A_f + (A_b - A_f) \frac{[L] + K_D + [R] - \sqrt{([L] + K_D + [R])^2 - 4[L][R]}}{2[L]}$$

where A_f and A_b are the measured anisotropy of free and bound, fluorescently labelled Orc6, respectively, $[L]$ is the concentration of Orc6 used in binding assays, $[R]$ is the concentration of ORC1–5 and K_D is the apparent dissociation constant. Owing to the limited sensitivity of our plate reader, binding experiments were performed at Orc6 concentrations above the K_D for wild-type Orc6 binding to ORC1–5; hence, curve fits are meant to aid visualization rather than to model the data explicitly. Mean and standard deviations of fluorescence anisotropy values were plotted as a function of ORC1–5 concentration after subtracting A_f values from respective mean fluorescence anisotropy values for visual comparison (Extended Data Fig. 4f).

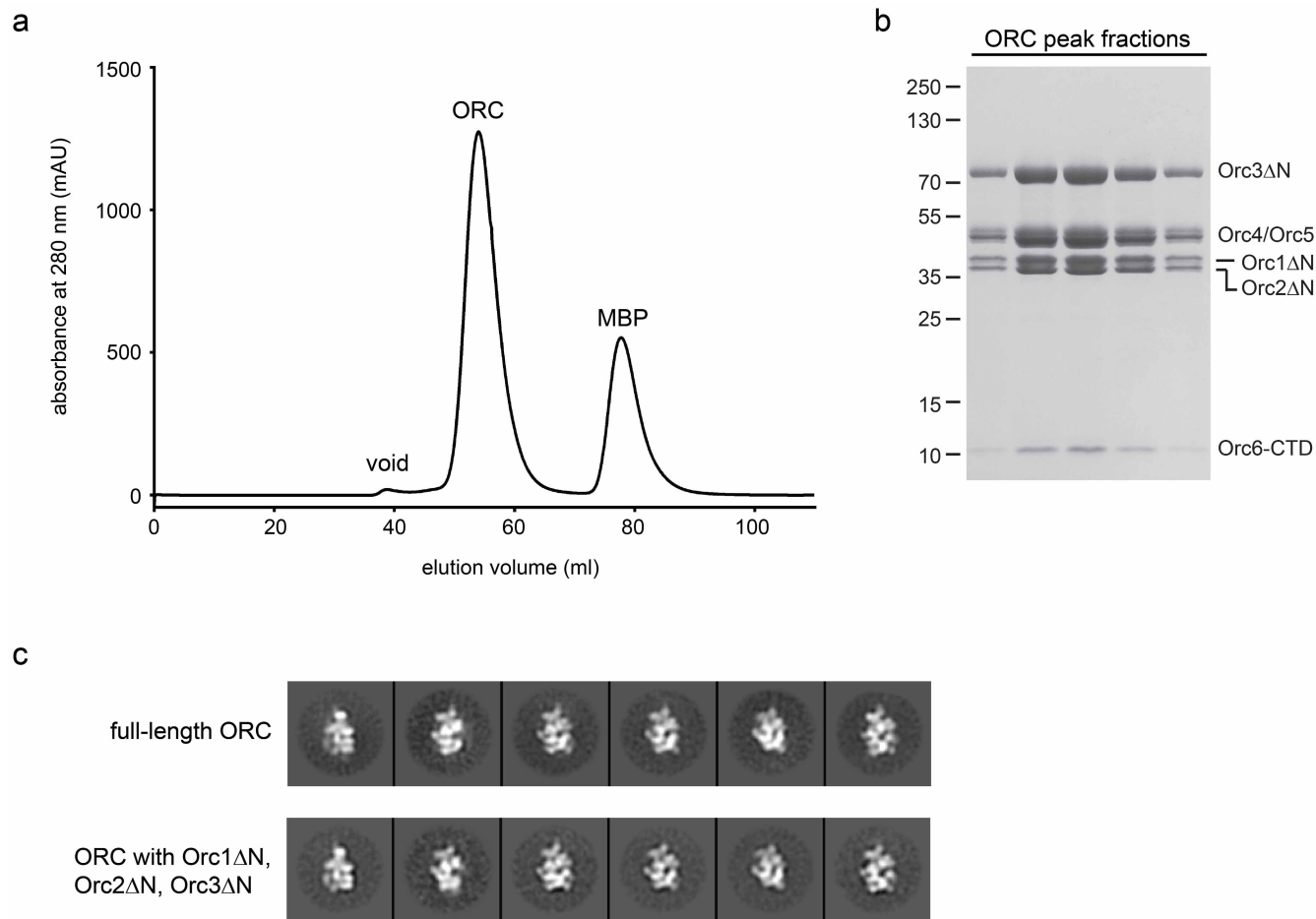
Crosslinking of Orc6 to ORC1–5. To independently check that the Orc6 C-terminal α -helix was oriented correctly with respect to Orc3, crosslinking studies were performed with either wild-type or an A236C Orc6-CTD mutant and untagged ORC1–5 using the homobifunctional maleimide crosslinker BMOE (bis(maleimido)ethane; Thermo Scientific) under non-reducing conditions. Ala236 of Orc6 was chosen for mutation because inspection of the structure revealed that the insert of Orc3 contained a nearby native cysteine (Cys372), which we reasoned could form a crosslink with an A236C Orc6 mutant if our build register were correct (Extended Data Fig. 4e). Before crosslinking, both wild-type Orc6-CTD and Orc6-CTD^{A236C} proteins, as well as untagged ORC1–5, were dialysed overnight into 50 mM PIPES (pH 7.5), 300 mM KCl and 10% glycerol to remove reducing agents. Binding of the mutant or wild-type Orc6 protein to ORC1–5 was performed in 50 μ l reactions containing 4 μ M of each protein in 50 mM PIPES (pH 7.5), 300 mM KCl and 10% glycerol for 30 min. BMOE was then added to reactions to a final concentration of 0.2 mM for 5 min, after which crosslinking reactions were stopped by adding SDS-polyacrylamide gel electrophoresis (SDS-PAGE) sample-loading dye containing β -mercaptoethanol (100 mM final concentration). Stopped reactions were analysed by SDS-PAGE and Coomassie staining.

To ensure that Orc6 still bound to ORC1–5 under non-reducing conditions, we performed binding reactions using ORC1–5 containing MBP-tagged Orc4 and either wild-type or A236C Orc6-CTD under non-reducing conditions as described for the crosslinking experiments, but instead subjected the reactions to pull-downs using amylose beads (New England Biolabs). Beads were washed three times with 1 ml of 50 mM PIPES (pH 7.5), 300 mM KCl and 10% glycerol, after which bound proteins were eluted with buffer containing 20 mM maltose and analysed by SDS-PAGE.

Electron microscopy. Four microlitres of a 30 nM *Drosophila* ORC solution containing N-terminally truncated Orc1, Orc2 and Orc3 subunits (in 20 mM Tris (pH 7.8), 125 mM potassium glutamate, 5 mM MgCl₂, 1 mM ATP γ S) were spotted onto glow-discharged, continuous-carbon film EM grids, stained with four drops of 2% uranyl formate for 10 s each and blotted. Grids were imaged in a Tecnai T12 BIOTWINN transmission electron microscope operated at 120 keV with a LaB₆ cathode as electron source. Data collection, image processing and two-dimensional classification were performed as described previously¹⁴, and two-dimensional class averages were compared with class averages of the full-length ATP γ S-ORC data set reported in ref. 14 (Extended Data Fig. 1c).

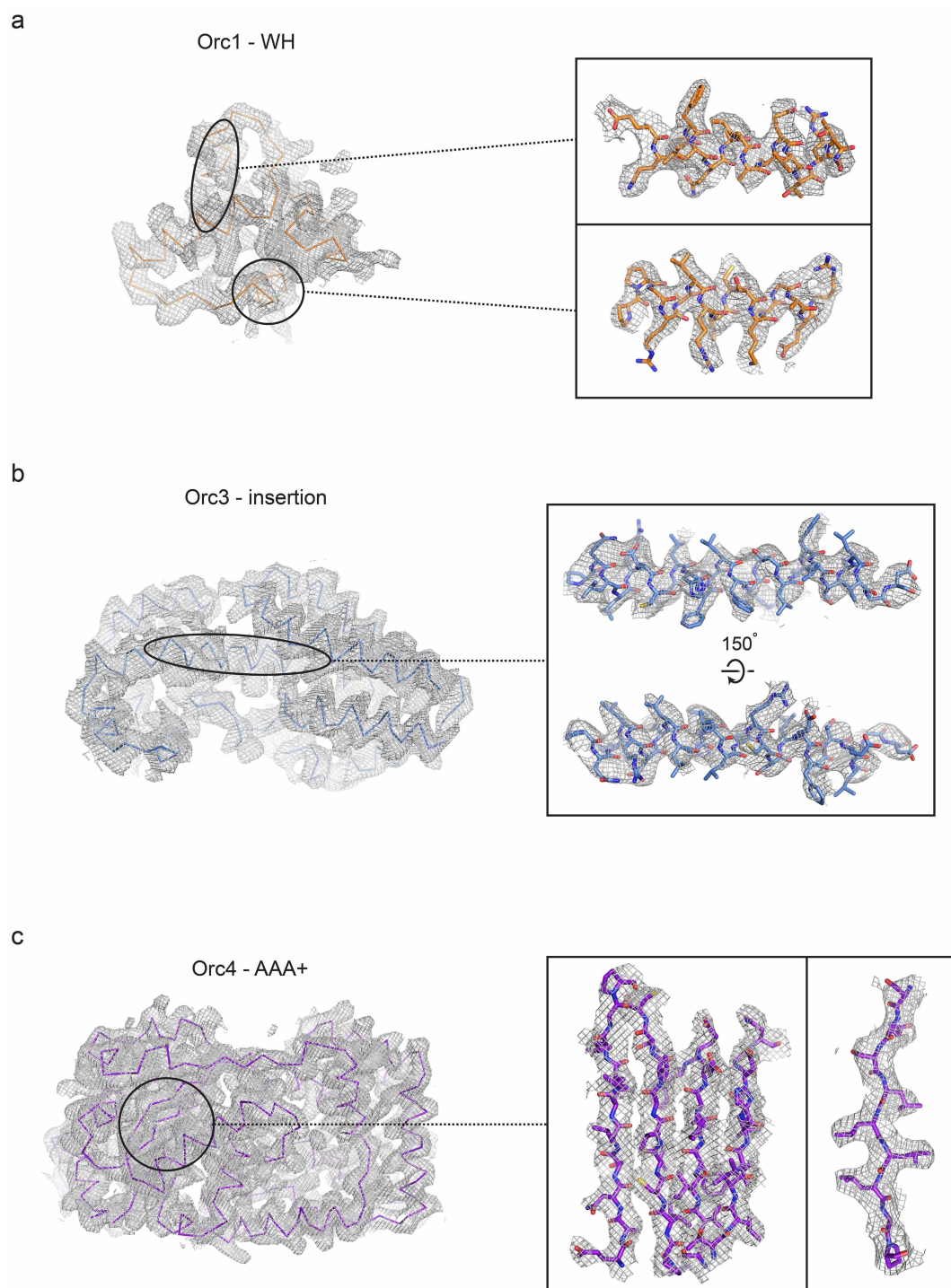
49. Parks, T. D. *et al.* Release of proteins and peptides from fusion proteins using a recombinant plant virus proteinase. *Anal. Biochem.* **216**, 413–417 (1994).
50. Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).
51. Kabsch, W. Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D* **66**, 133–144 (2010).
52. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).
53. Sheldrick, G. M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D* **66**, 479–485 (2010).
54. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
55. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
56. Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).
57. Vornrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230 (2007).
58. Bricogne, G. *et al.* Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr. D* **59**, 2023–2030 (2003).
59. Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallogr. D* **66**, 22–25 (2010).

60. Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157**, 281–287 (2007).
61. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
62. Kelley, L. A. & Sternberg, M. J. Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols* **4**, 363–371 (2009).
63. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
64. Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).
65. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
66. Katoh, K. & Toh, H. Recent developments in the MAFFT multiple sequence alignment program. *Brief. Bioinform.* **9**, 286–298 (2008).
67. Ashkenazy, H. *et al.* ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38**, W529–W533 (2010).
68. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

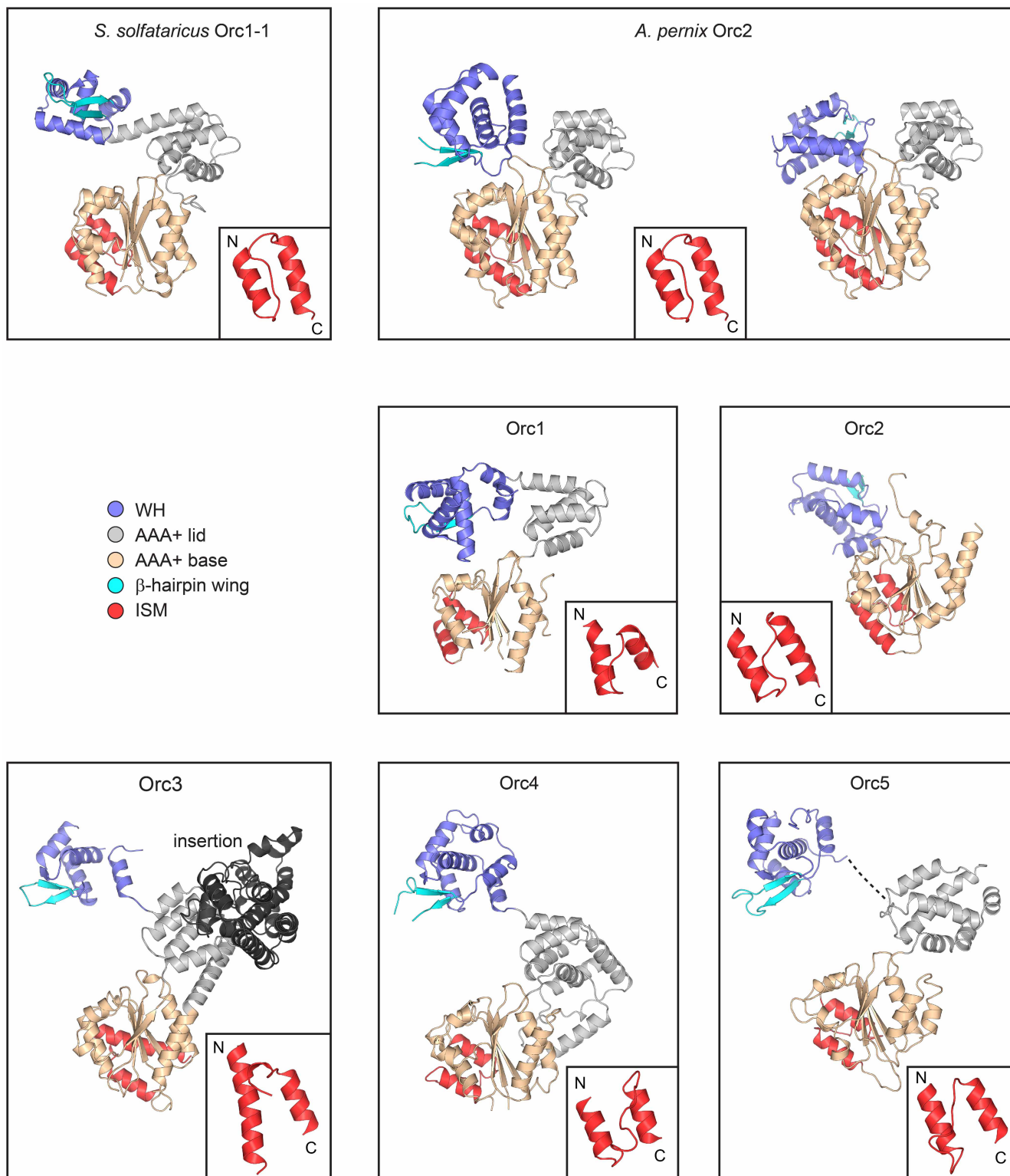


Extended Data Figure 1 | Deletion of variable N-terminal extensions in Orc1, Orc2 and Orc3 alters neither ORC stability nor overall ORC architecture. **a**, Gel-filtration chromatography trace of the ORC core used for crystallography, with **(b)** SDS-PAGE of respective ORC peak fractions from **a**, indicates the formation of a stable hexameric complex. **c**, Full-length ORC and ORC containing N-terminal truncations display similar structural features

in two-dimensional EM class averages. Both complexes were imaged by negative-stain EM in the presence of ATP γ S. Note that although class averages from ORC with truncated Orc1–3 subunits contain full-length Orc6, Orc6 is not visible owing to its flexible nature¹⁴. Class averages for full-length ORC are derived from a data set used in ref. 14.

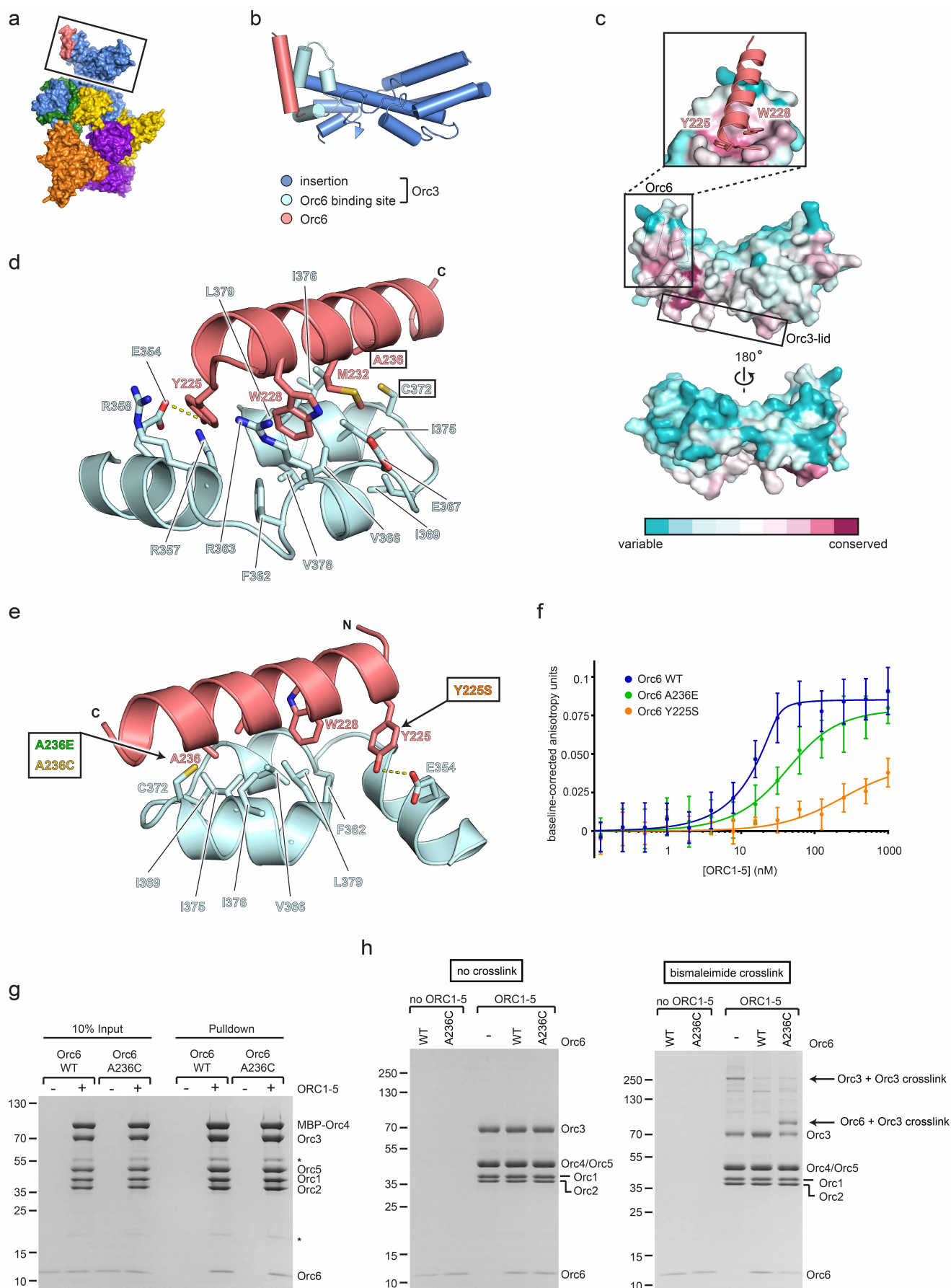


Extended Data Figure 2 | Experimental electron density contoured at 1σ for different regions of the ORC structure. The Orc1 WH domain is shown in **a**, the Orc3 insertion in **b** and the Orc4 AAA+ domain in **c**.



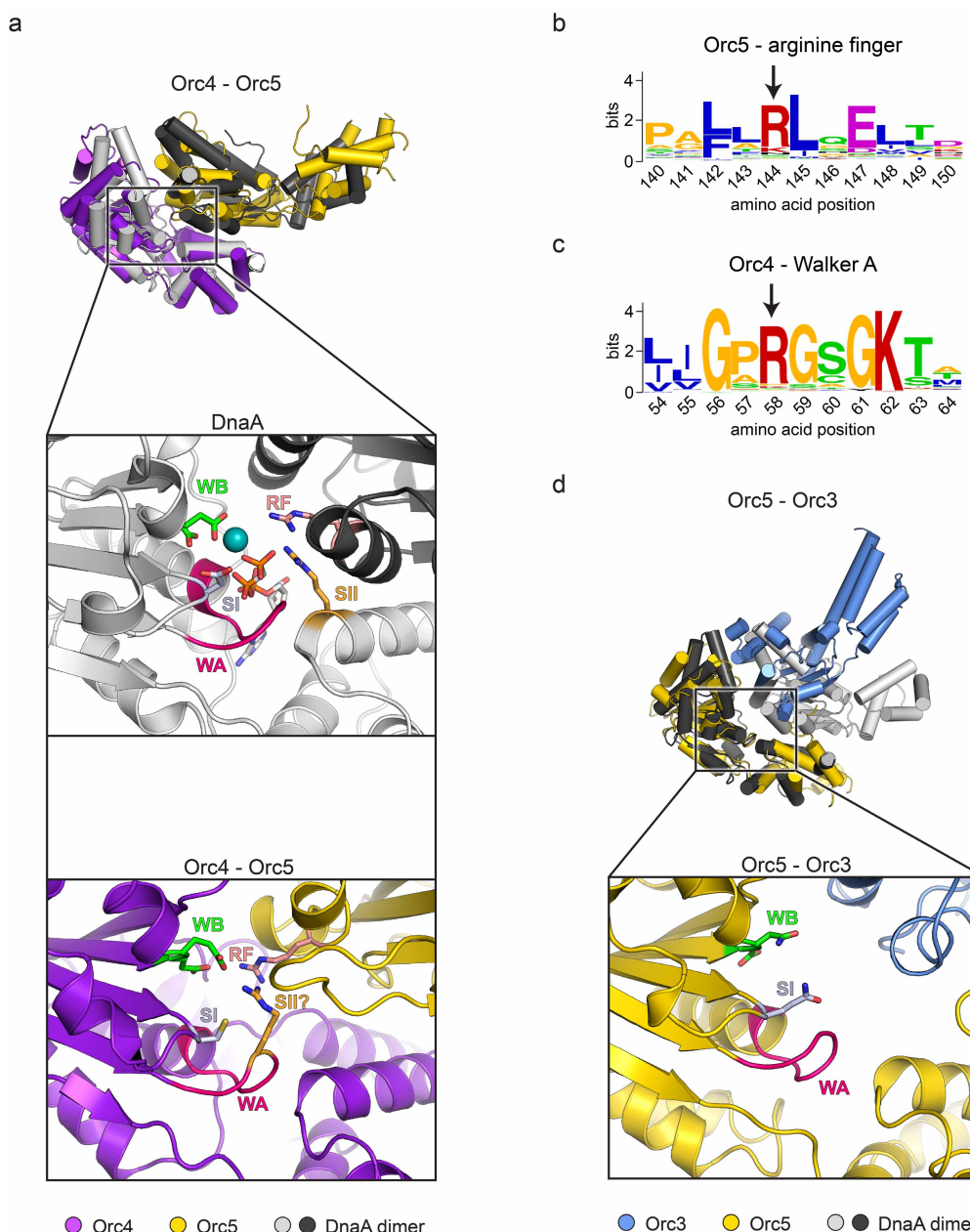
Extended Data Figure 3 | Comparison of individual ORC subunits and archaeal Orc structures. Structures of individual ORC subunits are compared to *S. solfataricus* Orc1-1 (PDB accession number 2QBY chain A¹⁶) and *A. pernix* Orc2 (PDB accession numbers 1W5S chain A (left) and 1W5T chain C (right)¹⁸). Different structural elements are coloured as indicated. The

initiator-specific motif (ISM) of the AAA+ ATPase fold is shown in the inset. No electron density was observed for the region linking the AAA+ and WH domains of Orc5 (indicated by a dashed line). The very N-terminal region of Orc2, which could only be built as stretches of polyalanine, is not shown.



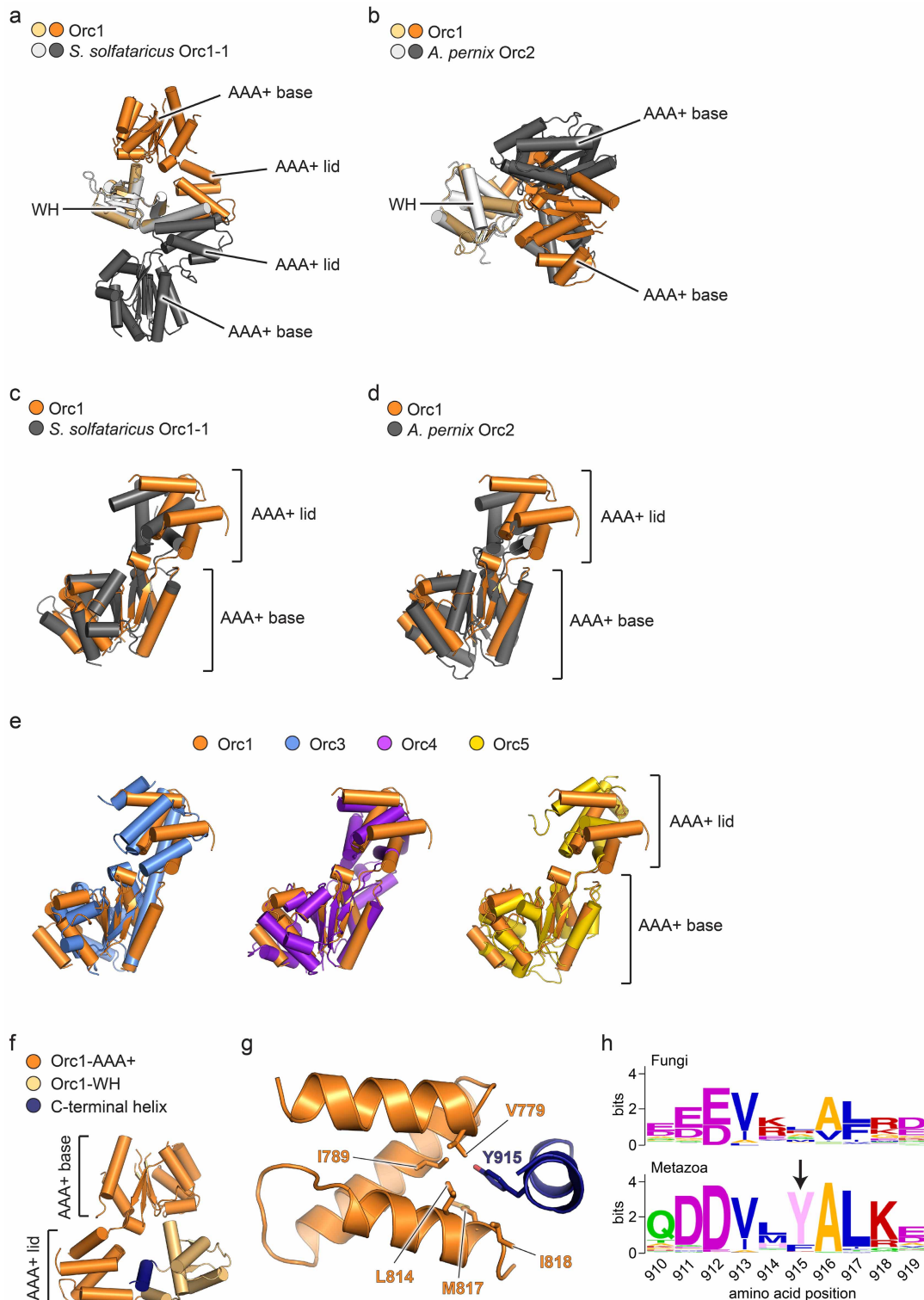
Extended Data Figure 4 | The Orc3 domain insertion forms a conserved, hydrophobic binding platform for Orc6. **a**, Surface representation of ORC. The Orc3 insertion, which extends from the Orc3 AAA + lid subdomain and interacts with the C-terminal helix of Orc6, is boxed. **b**, Secondary structure representation of the boxed region shown in **a**. The Orc3 insertion forms a bi-lobed, α -helical fold, three helices of which create a binding site for Orc6. **c**, Surface conservation of the Orc3 insertion. Conserved Orc3 residues cluster in the region that interacts with the Orc3 lid and in the Orc6 binding pocket. The latter region contacts highly conserved residues in Orc6 (Y225 and W228). **d**, Close-up view of Orc3•Orc6 interactions, showing a primarily hydrophobic binding site in Orc3 for Orc6 residues (Y225, W228, M232, A236). Y225, which in *Drosophila* Orc6 is equivalent to an amino acid altered in a subset of patients with Meier–Gorlin syndrome, appears positioned within hydrogen-bonding distance of E354 in Orc3 (dashed line). Colours are as in **b**. **e–h**, Biochemical validation of the binding register for *Drosophila* Orc6. **e**, Close-up of the Orc6•Orc3 interface. Orc6–Ala236 faces a hydrophobic surface formed by Orc3 residues and is also in close proximity to a natural cysteine in Orc3 (Cys372). To validate the register of the short C-terminal Orc6 helix and the Orc6•Orc3 interface, we mutated Orc6–Ala236 to either glutamate, which we hypothesized would impede binding to ORC1–5 because of clashes with hydrophobic residues in Orc3, or to cysteine, which we presumed would not affect Orc3 binding but would allow site-specific crosslinking to Orc3–Cys372. **f**, Orc6^{A236E} has a reduced affinity for the ORC1–5 complex. The C-terminal domains (CTDs) of wild-type (WT) Orc6, Orc6^{A236E} or the Meier–Gorlin syndrome equivalent Orc6^{Y225S} were each N-terminally labelled with Alexa Fluor 488 and tested for ORC1–5 binding using fluorescence anisotropy. As shown previously¹⁴, the C-terminal domain

of Orc6 binds ORC1–5 with low nanomolar affinity, whereas the Y225S mutation strongly reduces binding. As predicted on the basis of the structure of the Orc6•Orc3 interface, the A236E mutation also reduces the affinity of the Orc6-CTD for ORC1–5. Mean and standard deviations from three (for Orc6^{Y225S} and Orc6^{A236E}) or six (for wild-type Orc6) independent experiments are shown. **g**, Orc6^{A236C} is able to bind to the ORC1–5 complex. Orc6-CTD^{WT} or Orc6-CTD^{A236C} were incubated with ORC1–5 (containing MBP-tagged Orc4) and subjected to pull-down experiments using amylose resin. Both Orc6-CTD^{WT} and Orc6-CTD^{A236C} co-purified with ORC1–5. The pull-down experiment was performed under non-reducing experimental conditions similar to the crosslinking experiment in **h**. Asterisks mark two likely proteolytic fragments of Orc3. **h**, The Orc6-CTD^{A236C} mutant, but not the wild-type Orc6-CTD, specifically crosslinks to Orc3 within the ORC1–5 complex. Orc6-CTD^{WT} or Orc6-CTD^{A236C}, either alone or in the presence of ORC1–5, was incubated with a bifunctional maleimide crosslinker and the proteins subsequently analysed by SDS–PAGE. In reactions containing ORC1–5 and Orc6-CTD^{A236C}, crosslinking gives rise to a novel band with higher molecular mass than Orc3; the appearance of this band correlates with a decrease in the amount of uncrosslinked Orc3 and Orc6-CTD, and does not appear with reactions containing ORC1–5 and wild-type Orc6-CTD, indicating that this species corresponds to an Orc3–Orc6 crosslink (a moderately strong higher molecular-mass band that appears in the absence of Orc6 probably corresponds to homotypic adducts between exposed cysteines in Orc3). These results are consistent with the structure, which places Orc6–Ala236 in close proximity to Orc3–Cys372. Note that ORC1–5 contained MBP-tagged Orc4 in **g** but that the tag was removed in **h**.



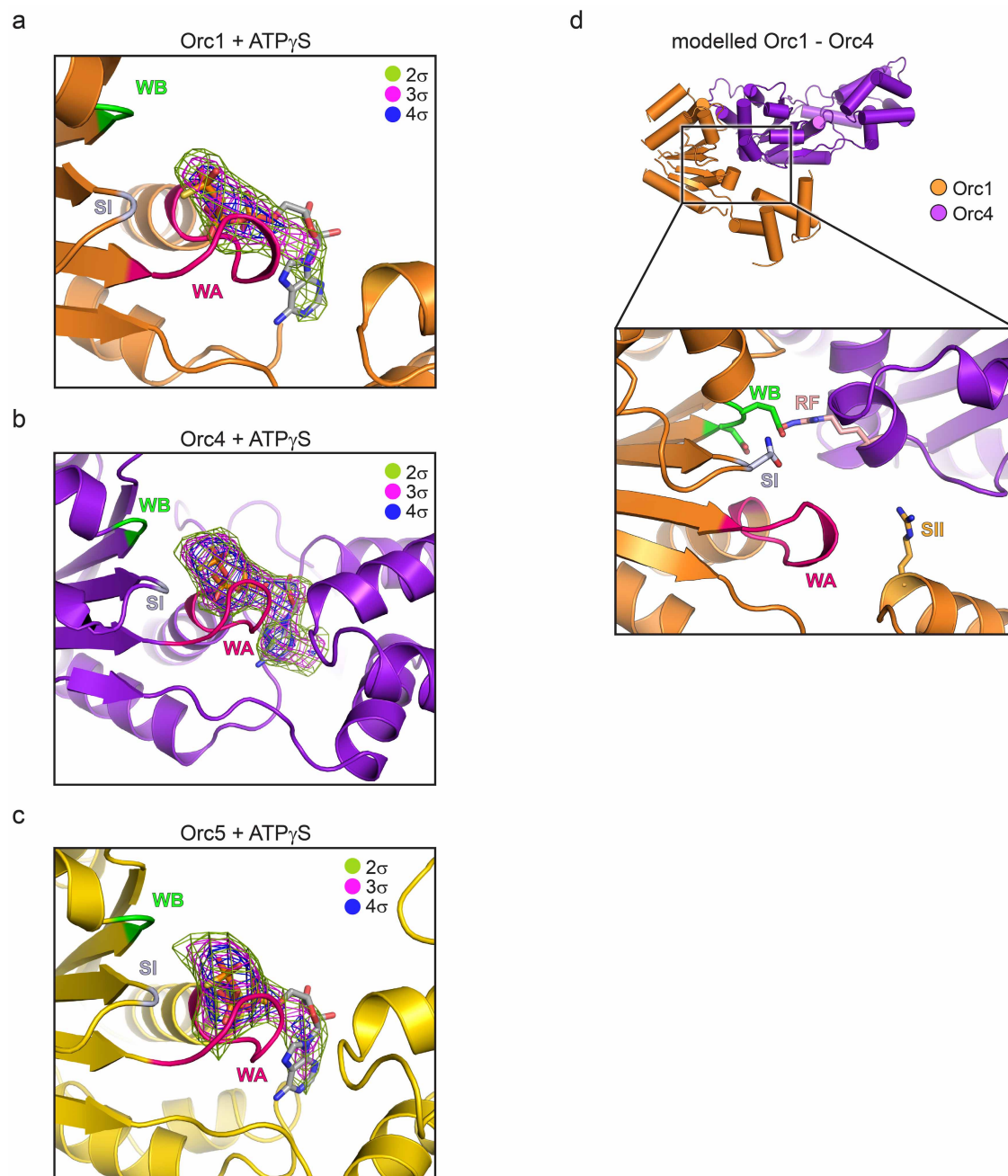
Extended Data Figure 5 | ATP-binding site configuration at the Orc4•Orc5 and Orc5•Orc3 interfaces. **a**, Inter-AAA+ interactions between Orc4 and Orc5 are similar to canonical AAA+ interactions between DnaA protomers (top panel, only Orc4 is used for superpositioning onto the left (light grey) AAA+ domain of an ATP-bound DnaA dimer, PDB accession number 2HCB²³). Close-up views of the nucleotide-binding site are shown for Orc4 (bottom panel) and for DnaA for comparison (middle panel). The resemblance of the Orc4 nucleotide-binding pocket to the active site of functional AAA+ ATPases is somewhat surprising considering that mutations in the active site of *Drosophila* and human Orc4 have no reported effect on the ATPase activity of ORC as measured *in vitro*^{28,29}, but may help explain why a *Drosophila* ORC mutant bearing a Walker A or B substitution in Orc4 exhibits modest DNA replication defects in extracts²⁸. **b**, The putative arginine finger in Orc5 is well conserved across homologues. A sequence logo of a multiple sequence alignment of the region containing the putative arginine finger (marked with an

arrow) in eukaryotic Orc5 protein sequences is shown. Amino-acid numberings correspond to the *Drosophila* Orc5 sequence. **c**, A potential Sensor II equivalent arginine (marked with an arrow) in the Orc4 Walker A motif is conserved in eukaryotic Orc4 homologues. A sequence logo of the Walker A motif from a multiple sequence alignment of eukaryotic Orc4 protein sequences is shown. Amino-acid positions are numbered as in *Drosophila* Orc4. **d**, Inter-AAA+ interactions between Orc5 and Orc3. The top panel shows a superposition derived from placing the AAA+ domain of Orc5 atop the AAA+ domain of the left (dark grey) protomer of an ATP-bound DnaA dimer; the bottom panel shows a close-up view of the nucleotide-binding site at the Orc5•Orc3 interface. Side chains of conserved residues known to be involved in nucleotide binding and hydrolysis in AAA+ ATPases are represented as sticks in both **a** and **d**. WA, Walker A; WB, Walker B; SI, Sensor I; SII, Sensor II; RF, arginine finger.



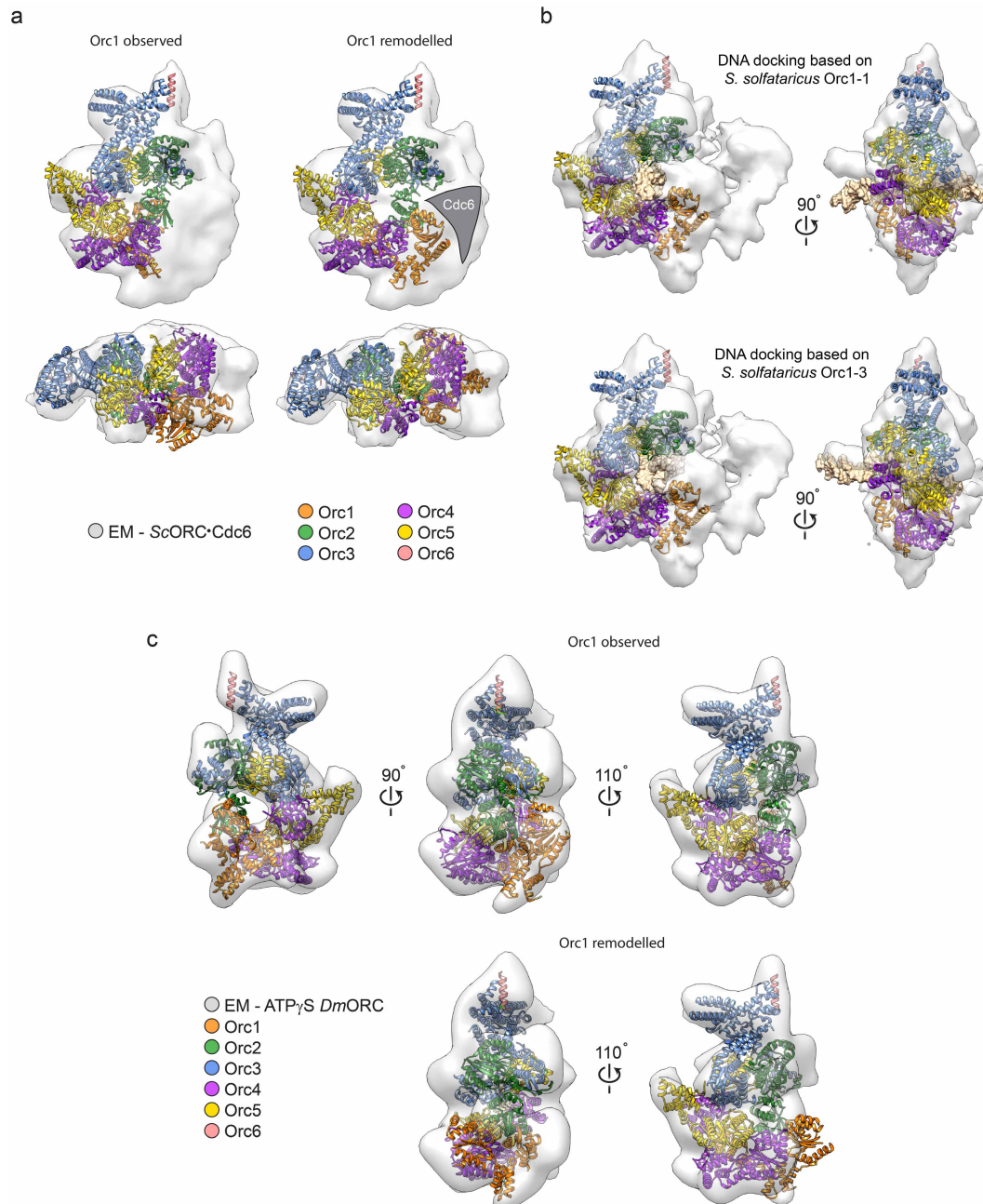
Extended Data Figure 6 | The conformation of Orc1 arises from a reorientation between its AAA+ and WH domains, not from changes within the AAA+ ATPase domain itself. **a**, Superpositioning of the WH domains of Orc1 and *S. solfataricus* Orc1-1 (PDB accession number 2QBY chain A¹⁶) reveals different conformations for both proteins, resulting from a large domain rotation of the Orc1 AAA+ domain around a pivot point in the linker preceding its WH domain. **b**, The Orc1 conformation is most similar to a state seen for *A. permix* Orc2 (PDB accession number 1W5T chain C¹⁸). The WH domains of both proteins were superposed as in **a**. **c–e**, Superposing the AAA+ base subdomains of Orc1 and *S. solfataricus* Orc1-1 (**c**, PDB accession number 2QBY chain A¹⁶), *A. permix* Orc2 (**d**, PDB accession number 1W5T chain C¹⁸) and Orc3, Orc4 or Orc5 (**e**) shows that the typical AAA+

configuration between the base and lid subdomains are maintained in Orc1. Only a slight opening of the nucleotide-binding cleft is observed in Orc1, which is probably caused by the absence of bound nucleotide. **f**, **g**, The most C-terminal α -helix of the Orc1 WH domain mediates interactions with the Orc1 lid subdomain. An overview of the interaction is shown in **f**, with a close-up view of contacts between a conserved tyrosine (Tyr915) in the C-terminal Orc1 helix and a hydrophobic pocket of the Orc1 lid depicted in **g**. **h**, The tyrosine in the C-terminal helix of Orc1 is well conserved across metazoan but not fungal Orc1 homologues. Alignments are shown as sequence logos. The numbering of amino acids is based on *Drosophila* Orc1, and the tyrosine is marked by an arrow.



Extended Data Figure 7 | Nucleotide binding by Orc1, Orc4 and Orc5. For **a–c**, molecular replacement with the *apo*-ORC model was used to phase diffraction data collected from an ORC-ATP γ S co-crystal. Positive $F_o - F_c$ difference density contoured at different sigma levels reveals clear features for nucleotide binding to the AAA+ domains of (a) Orc1, (b) Orc4 and (c) Orc5. ATP γ S is docked into the difference density for reference; owing to the moderate (4.0 Å) resolution of the data, this structure was not refined.

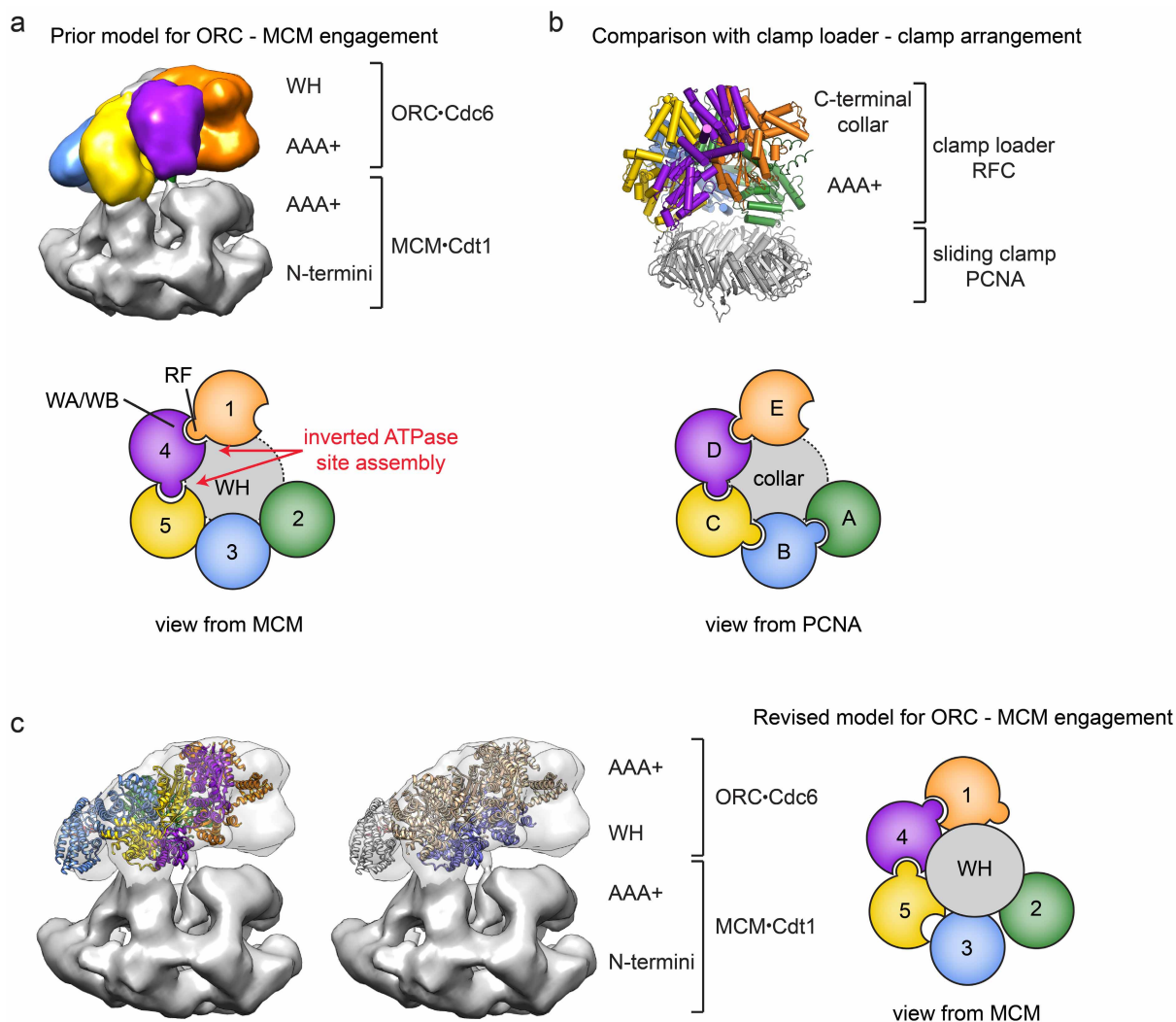
d, Modelling of canonical AAA+ interactions between Orc1 and Orc4, generated using the Orc4•Orc5 interaction as a reference. Upper panel: structural overview of modelled AAA+ domain positioning between Orc1 and Orc4. Lower panel: close-up of the modelled Orc1•Orc4 ATPase site. Side chains (taken from their place in the *apo*-ORC model as a reference) are shown for conserved catalytically important residues.



Extended Data Figure 8 | Comparison of crystallographic and EM models.

Docking of the observed and remodelled ORC structures into the cryo-EM density of *S. cerevisiae* and *Drosophila* ORC indicates that the ATPase domain of Orc1 is repositioned into a canonical AAA+/AAA+ interaction with Orc4 when Cdc6 is present, and supports a model where DNA passes through the central channel in ORC. **a**, The three-dimensional EM volume for *S. cerevisiae* ORC (as present in a complex with Cdc6, Cdt1 and MCM2–7 and assembled in the presence of DNA, EMD-5625 (ref. 24)) contains Orc1 in the activated conformation. ORC with Orc1 in the autoinhibited conformation (left panel, as observed in the crystal structure) and remodelled conformation (right panel, remodelled) were docked into the ORC•Cdc6•Cdt1•Mcm2–7 cryo-EM map (only the density for ORC•Cdc6 is shown). The ORC•Cdc6 EM density readily accommodates Orc1 in the activated conformation, but not in its autoinhibited state. The EM density corresponding to Cdc6 is indicated in the right panel. **b**, DNA passes through the central ORC channel in the DNA•ORC•Cdc6 complex. ORC (with Orc1 in the remodelled conformation) was first docked into the cryo-EM map derived from a DNA•ORC•Cdc6 complex (EMD-5381 (ref. 23)). The AAA+ domain and its associated DNA from either *S. solfataricus* Orc1-1 or Orc1-3 (PDB accession number 2QBY¹⁶) were then superposed using the AAA+ domain of Orc4 as a guide. Both dockings indicate that a region of density previously assigned to the Orc6

subunit²³ actually corresponds to the DNA duplex. Although superpositioning of the AAA+ domain of *S. solfataricus* Orc1-3 onto Orc4 better positions duplex DNA in the observed EM density (than does the comparable exercise using the Orc1-1•DNA complex), the curvature of the DNA (as present in the Orc1-3•DNA co-crystal structure) results in a greater number of clashes between DNA and ORC subunits. Nevertheless, both docking scenarios are consistent with a DNA binding mode of ORC where DNA runs through the central channel. Note that the handedness of the EM map (EMD-5381 (ref. 23)) has been corrected in this figure because it has been reported that the original handedness was inverted²⁴. For clarity, the WH domain of Orc2 is omitted from the remodelled ORC structure in **b**, **c**. The autoinhibited ORC conformation observed in the crystal is, unlike the remodelled Orc1 configuration, similar to the ORC conformation observed in *Drosophila* ORC EM reconstructions. Docking of the ORC crystal structure (top panels) or the remodelled activated ORC structure (bottom panels) into a prior three-dimensional EM reconstruction of *Drosophila* ORC (EMD-2479)¹⁴ reveals excellent agreement between EM and crystal structures, but not between EM and modelled activated ORC structures. The poor fit of the remodelled Orc1 conformation into the EM density suggests that the EM structure represents the autoinhibited state of ORC as seen in the crystal, indicating it is the predominant state in solution. See also Supplementary Video 2.



Extended Data Figure 9 | Docking of the ORC structure into the cryo-EM structure of an *S. cerevisiae* replication initiation intermediate indicates that ORC recruits the MCM2–7 complex by binding to the ORC WH domains. A prior model for ORC·MCM2–7 engagement²⁴, proposed from an ORC·Cdc6·Cdt1·MCM2–7 cryo-EM structure generated in the presence of DNA (shown in **a**, EMD-5625 (ref. 24)), used the crystal structure of replication factor C (RFC) bound to the sliding clamp PCNA (shown in **b**, PDB accession number 1SXJ³¹) to suggest that the AAA+ domains of ORC engage the MCM2–7·Cdt1 complex. However, using the handedness of the EM volume as reported²⁴, this organization of ORC subunits leads to an inverted ATPase site assembly, requiring that the Orc4 arginine finger (which is known to

stimulate Orc1 ATP hydrolysis³³) points towards the Orc5 nucleotide-binding site rather than the appropriate Orc1 active site. Schematics for the ATP site assemblies of ORC and RFC derived from these structures are shown in the lower panels in **a** and **b**. The location of the WH domain collar of ORC and the C-terminal collar of RFC is indicated by a grey circle. **c**, Docking of the ORC crystal structure (with Orc1 in its remodelled or ‘activated’ conformation) into the cryo-EM map shown in **a** reveals that the WH domains of ORC face an MCM2–7 complex. This switched polarity of WH domains and AAA+ domains in the EM map corrects the ATPase site assembly and is schematized in the right panel.

Extended Data Table 1 | Summary of data collection, phasing and refinement statistics

	Native	GdCl ₃ soak	Native - ATPyS
Data collection:			
Beamline	APS 23ID-B	APS 23ID-B	NSLS X25
Wavelength (Å)	1.0332	1.71083	1.1
Resolution range (Å)	48.8-3.5 (3.59-3.5)	50.23-3.96 (4.13-3.96)	58.11-4.0 (4.16-4.0)
Space group	I222	I222	I222
Unit cell dimensions			
a, b, c (Å)	145.5, 259.0, 257.0	145.3, 257.7, 256.1	144.2, 255.3, 261.1
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Reflections			
Total	518,695 (38,421)	2,359,387 (209,066)	554,356 (62,459)
Unique	61,418 (4,491)	40,955 (3,602)	41,025 (4,568)
Completeness (%)	100.0 (100.0)	97.3 (76.2)	100.0 (100.0)
Multiplicity	8.4 (8.6)	57.6 (58.0)	13.5 (13.7)
Mean I/σ	7.7 (1.1)	23.4 (1.9)	16.6 (1.3)
CC _{1/2}	0.995 (0.427)	0.999 (0.745)	0.999 (0.678)
R _{merge}	0.214 (2.181)	0.355 (3.857)	0.095 (2.174)
R _{pim}	0.078 (0.788)	0.047 (0.509)	0.027 (0.606)
Wilson B factor	118.6	165	197.8
Phasing:			
FOM before density modification		0.27	
FOM after density modification		0.64	
Refinement:			
R _{work} /R _{free}	0.2243/0.2581		
Number of atoms			
Protein	16336		
Ligands	2		
B factors			
Mean	118.1		
Macromolecule	118.1		
Ligands	109.2		
Root mean square deviation			
Bond lengths (Å)	0.006		
Bond angles (°)	0.705		
Ramachandran plot			
% favored	92.9		
% allowed	7.1		
% outliers	0.0		
Molprobrity			
Clashscore	4.53		

Highest resolution shell is shown in parenthesis.

A dusty, normal galaxy in the epoch of reionization

Darach Watson¹, Lise Christensen¹, Kirsten Kraiberg Knudsen², Johan Richard³, Anna Gallazzi^{1,4} & Michał Jerzy Michałowski⁵

Candidates for the modest galaxies that formed most of the stars in the early Universe, at redshifts $z > 7$, have been found in large numbers with extremely deep restframe-ultraviolet imaging¹. But it has proved difficult for existing spectrographs to characterize them using their ultraviolet light^{2–4}. The detailed properties of these galaxies could be measured from dust and cool gas emission at far-infrared wavelengths if the galaxies have become sufficiently enriched in dust and metals. So far, however, the most distant galaxy discovered via its ultraviolet emission and subsequently detected in dust emission is only at $z = 3.2$ (ref. 5), and recent results have cast doubt on whether dust and molecules can be found in typical galaxies at $z \geq 7^{6–8}$. Here we report thermal dust emission from an archetypal early Universe star-forming galaxy, A1689-zD1. We detect its stellar continuum in spectroscopy and determine its redshift to be $z = 7.5 \pm 0.2$ from a spectroscopic detection of the Lyman- α break. A1689-zD1 is representative of the star-forming population during the epoch of reionization⁹, with a total star-formation rate of about 12 solar masses per year. The galaxy is highly evolved: it has a large stellar mass and is heavily enriched in dust, with a dust-to-gas ratio close to that of the Milky Way. Dusty, evolved galaxies are thus present among the fainter star-forming population at $z > 7$.

As part of a programme to investigate galaxies at $z > 7$ with the X-shooter spectrograph on the Very Large Telescope, we observed the candidate high-redshift galaxy, A1689-zD1, behind the lensing galaxy cluster Abell 1689 (Fig. 1). The source was originally identified¹⁰ as a candidate $z > 7$ system from deep imaging with the Hubble and Spitzer space telescopes, with photometry fitting suggesting that it is at $z = 7.6 \pm 0.4$. The galaxy is gravitationally magnified by a factor of 9.3 by the galaxy cluster¹⁰. Although it is intrinsically faint, because of the gravitational amplification, it is one of the brightest candidate $z > 7$ galaxies known. The X-shooter observations were carried out on several nights between March 2010 and March 2012 with a total time of 16 h on target.

The galaxy continuum is detected and can be seen in the binned spectrum (Fig. 2). The Ly α cutoff is at $1,035 \pm 24$ nm and defines the redshift to be $z = 7.5 \pm 0.2$. It is thus one of the most distant galaxies known so far to be confirmed via spectroscopy, and the only galaxy at $z > 7$ where the redshift is determined from spectroscopy of its stellar continuum. The spectral slope is blue; using a power-law fit $F_\lambda \propto \lambda^{-\beta}$, $\beta = 2.0 \pm 0.1$, where λ is the wavelength and F_λ is the flux per unit wavelength. The flux break is sharp, and greater than a factor of ten in depth. In addition, no line emission is detected, ruling out a different redshift solution for the galaxy. Line emission is excluded to lensing-corrected

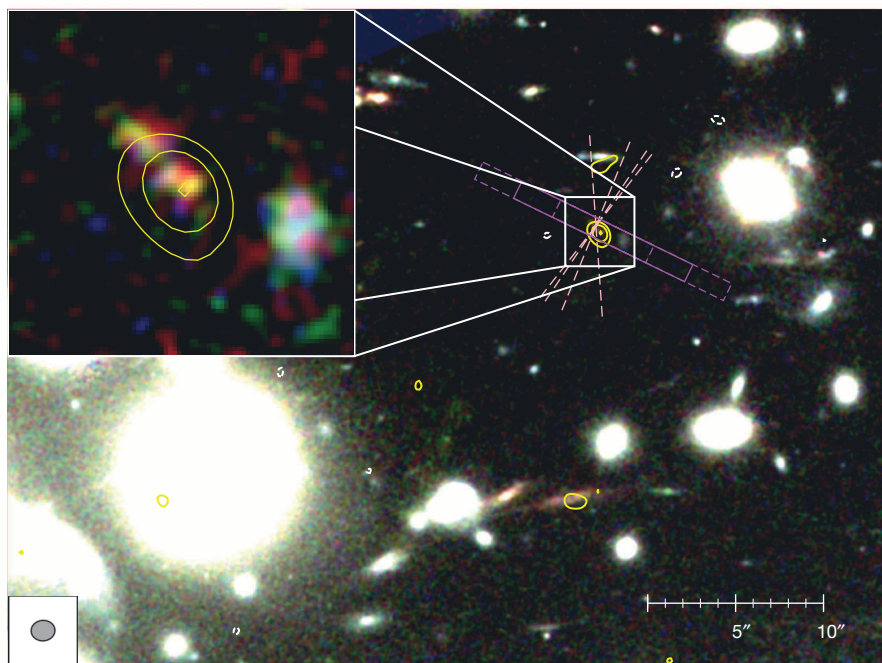


Figure 1 | The gravitationally lensing galaxy cluster Abell 1689. The colour image is composed with Hubble Space Telescope filters: F105W (blue), F125W (green) and F160W (red). The zoomed box ($4'' \times 4''$) shows A1689-zD1. Contours indicate far-infrared dust emission detected by ALMA at 3σ , 4σ , and 5σ local rms (yellow, positive; white, negative). The ALMA beam

($1.36'' \times 1.15''$) is shown, bottom left. Images and noise maps were primary-beam corrected before making the signal-to-noise ratio (SNR) maps. Slit positions for the first set of X-shooter spectroscopy are overlaid in magenta (dashed boxes indicate the dither), while the parallactic angle was used in the remaining observations (pink dashed lines).

¹Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 32, København Ø, 2100, Denmark. ²Department of Earth and Space Sciences, Chalmers University of Technology, Onsala Space Observatory, SE-439 92 Onsala, Sweden. ³Centre de Recherche Astrophysique de Lyon, Observatoire de Lyon, Université Lyon 1, 9 Avenue Charles André, 69561 Saint Genis Laval Cedex, France. ⁴Istituto Nazionale di Astrofisica-Osservatorio Astronomico di Arcetri, Largo Enrico Fermi 5, 50125 Firenze, Italy. ⁵The Scottish Universities Physics Alliance, Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK.

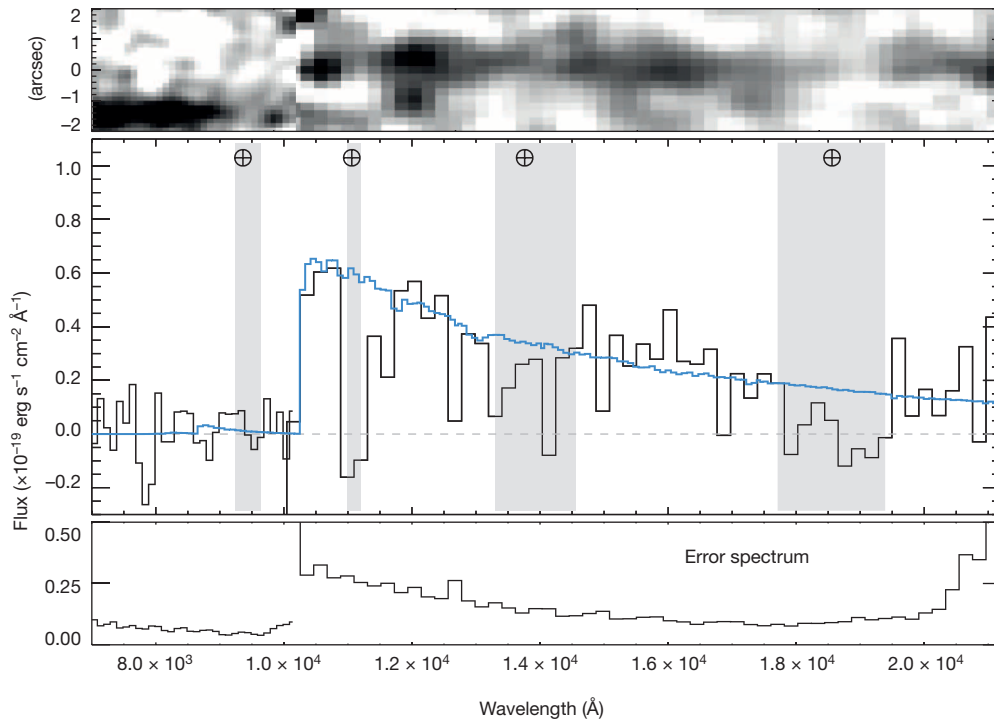


Figure 2 | Spectrum of A1689-zD1. The binned one-dimensional (middle panel) and two-dimensional (upper panel; wavelength versus distance along the slit) spectra are shown, with the 68% confidence uncertainty for the one-dimensional spectrum in the bottom panel. The redshift $z = 7.5$ is determined from the Ly α break at 1,035 nm. Sky absorption (grey bands) and

the best-fit SED (blue line) are shown. The Ly α break is close to the spectrograph's near-infrared (NIR)/visual (VIS) arm split; however, the break is clearly detected in the NIR arm alone. A nearby galaxy ($z \approx 2$) visible in the bottom part of the two-dimensional spectrum is detected in both the VIS and the NIR arms.

depths of $3 \times 10^{-19} \text{ erg cm}^{-2} \text{ s}^{-1}$ (3σ) in the absence of sky emission lines, making this by far the deepest intrinsic spectrum published of an object from the epoch of reionization, highlighting the difficulty of obtaining ultraviolet redshifts for objects at this epoch that are not strongly dominated by emission lines. The restframe equivalent width limits are $< 4 \text{ \AA}$ for both Ly α and C III] 1,909 \AA . Our search space for Ly α is largely free of sky emission lines; they cover 16% of the range.

Fits to the galaxy's spectral energy distribution (SED) yield a lensing-corrected stellar mass of 1.7×10^9 solar masses (M_{\odot}) (that is, $\log(M_{\star}/M_{\odot}) = 9.23^{+0.15}_{-0.16}$), with a best-fitting stellar age of 80 million years (that is, a light-weighted age t of $\log[t \text{ (yr)}] = 7.91^{+0.26}_{-0.24}$). The lensing-corrected ultraviolet luminosity is about $1.8 \times 10^{10} L_{\odot}$, where L_{\odot} is the solar luminosity, resulting in a star-formation rate (SFR) estimate of $2.7 \pm 0.3 M_{\odot} \text{ yr}^{-1}$ based on the ultraviolet emission and uncorrected for dust extinction, for a Chabrier initial mass function¹¹. A1689-zD1 is thus fainter than the turnover luminosity, L^* , in the galaxy luminosity

function at this redshift, meaning that it is among the faint galaxies that dominate star formation at this epoch⁹.

Mosaic observations of the lensing cluster were obtained with the Atacama Large Millimetre Array (ALMA) in Cycles 0 and 1 with the receivers tuned to four 3.8 GHz frequency bands between 211 GHz and 241 GHz. A1689-zD1 is located towards the northern edge of the mosaic and is detected at 5.0σ with an observed flux of $0.61 \pm 0.12 \text{ mJy}$ in the combined image and at $2.4\text{--}3.1\sigma$ significance in each of the three individual observations (Fig. 3). A1689-zD1 is located within the primary beam's full-width at half-maximum (FWHM) of one pointing and the sensitivity (root mean square, rms) around its position is 0.12 mJy per beam, 42% of the sensitivity of the deepest part of the mosaic. The source is the brightest in the mosaic area of 5 square arcminutes. It coincides with the ultraviolet position of A1689-zD1 and is $1.5''$ away from the next-nearest object in the Hubble Space Telescope image, which is not detected in the ALMA map. No line emission is convincingly detected

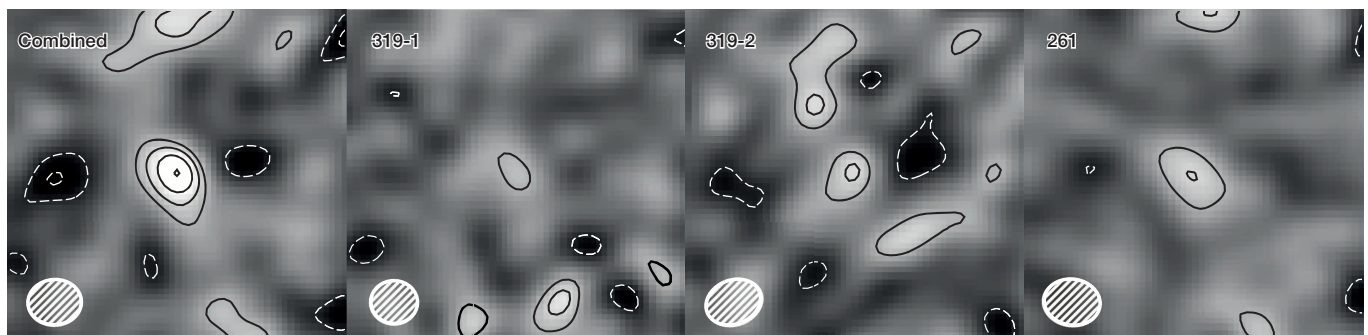


Figure 3 | ALMA SNR maps of A1689-zD1. Contours are SNR = 5, 4, 3, 2 (black, solid), -3, -2 (white, dashed). Images and noise maps were primary-beam corrected before making SNR maps. Beam sizes are shown at the bottom left of each panel. Panels are $8'' \times 8''$. The panels show from left to right: the combined data, the two tunings of observation 2011.0.00319.S and

observation 2012.1.00261.S. A1689-zD1 is detected from left to right, at 5.0σ , 2.4σ , 3.1σ , and 3.0σ . Natural weighting was used and the visibilities were tapered with a $1''$ circular Gaussian kernel, resulting in beams of $1.36'' \times 1.15''$, $1.19'' \times 1.09''$, $1.43'' \times 1.12''$, $1.43'' \times 1.17''$ from left to right.

Table 1 | Comparison of A1689-zD1 to other high-*z* star-forming galaxies

Galaxy name	Redshift, <i>z</i>	Stellar mass, <i>M</i> * ($10^9 M_\odot$)	SFR _{UV} ($M_\odot \text{ yr}^{-1}$)	SFR _{Lyα} ($M_\odot \text{ yr}^{-1}$)	SFR _{IR} ($M_\odot \text{ yr}^{-1}$)	Dust mass, <i>M</i> _D ($10^7 M_\odot$)
HFLS3 (ref. 21)	6.34	50^{+100}_{-30}	$1.3 \pm 0.4^*$	—	$1,300_{-520}^\dagger$	30_{-10}^\dagger
HCM6A (ref. 22)	6.56	—	9 ± 2	2	<28 (ref. 26)	<10 (ref. 26)
Himiko (ref. 8)	6.60	15 ± 2	30 ± 2	35 ± 1	<8	<4.72 (ref. 26)
A1703-zD1 (ref. 23)	6.8	0.7–1.5	7.3 ± 0.3	—	<16 (ref. 26)	<5.7 (ref. 26)
IOK-1 (ref. 24)	6.96	<40	23.9 ± 1.4 (ref. 27)	10 ± 2	<10 (ref. 28)	<6.4 (ref. 28)
z8-GND-5296 (ref. 2)	7.51	$1^{+0.2}_{-0.1}$	330^{+710}_{-10}	—	<127 (ref. 26)	<50 (ref. 26)
HG090423 (ref. 25)	8.2	<0.05 (ref. 29)	<0.38 (ref. 30)	—	<5 (ref. 29)	<2 ‡
A1689-zD1	7.5	$1.7^{+0.7}_{-0.5}$	2.7 ± 0.3	<0.7	9^{+4}_{-2}	4^{+4}_{-2}

The SFRs are derived from extinction-uncorrected ultraviolet emission, Ly α emission and far-infrared emission, respectively.

* Derived from the Hubble Space Telescope F160W photometry and corrected for lensing.

† 95% lower bound only.

‡ Assuming the same dust parameters assumed for A1689-zD1.

in the ALMA spectral data, which covers about half of the frequency range in which we could expect to find the [C II] 158 μm line, based on the range of redshifts allowed by the X-shooter Ly α break detection (see Methods). The derived total infrared luminosity L_{TIR} corrected for lensing and for cosmic microwave background (CMB) effects¹² is $6.2 \times 10^{10} L_\odot$, corresponding to a SFR of about $9 M_\odot \text{ yr}^{-1}$, with a dust mass in the galaxy of $4 \times 10^7 M_\odot$, assuming a dust temperature of 35 K. The dust mass we determine is conservative and is unlikely to be lower than about $2 \times 10^7 M_\odot$. It could in principle be larger by an order of magnitude, owing to temperature and geometry effects (see Methods). Table 1 presents a comparison of the properties of A1689-zD1 with other $z > 6.5$ galaxies with sensitive millimetre-wavelength observations. Most of the sources observed so far are more ultraviolet-luminous than A1689-zD1 and have upper limits somewhat above the dust mass found here.

Far-infrared emission requires the production of metals in a galaxy, whether in the solid phase as dust, or as ionized gas. To be detected, such a galaxy must have enriched its interstellar media with metals and dust. The metals are primarily produced and distributed via supernova explosions and so metal enrichment happens concurrently with massive star formation, but the site of dust production is less certain. However, the mechanism must be very rapid¹³ and these observations of A1689-zD1 place the strongest direct constraints so far on the rapidity of dust enrichment, occurring within only 500 million years of the beginning of star formation in the Universe.

We can derive an approximate gas mass for this galaxy by inverting the Schmidt–Kennicutt law based on its size and SFR¹⁴ and we deduce a gas mass of $2^{+2}_{-1} \times 10^9 M_\odot$. This gives a dust-to-gas mass ratio of about 17×10^{-3} . And while the uncertainty on the gas and dust masses is large—approximately 0.5 dex, dominated by the scatter in the Schmidt–Kennicutt law and the unknown dust temperature, where the two values are linked through the SFR—the dust-to-gas mass ratio is nevertheless high for this redshift, between a half and a few times the Milky Way value¹⁵. For a constant dust-to-metals ratio¹⁶, this also suggests a high metallicity with a similar uncertainty. Such a high metallicity is consistent with the fundamental mass–metallicity–SFR relation for star-forming galaxies¹⁷, from which we determine an expected elemental abundance of about half the solar value. Instead of a young, dust-poor galaxy, these measurements suggest an evolved system. Finally, the deep upper limit on the C III] 1,909 Å emission line, of $<4 \text{ Å}$ restframe equivalent width, is unusual for line-emitting galaxies at these redshifts¹⁸. From galaxies at lower redshift, we might expect an equivalent width as high as 30 Å for this SFR for a young galaxy¹⁹. And while the lack of Ly α emission in this galaxy could be explained by absorption in the intergalactic medium, the absence of C III] emission cannot, and is consistent with a more evolved galaxy. The surface area we are observing in ultraviolet emission is only approximately 1.5 kpc² (lensing-corrected size), so the SFR per unit area is high, approximately $8 M_\odot \text{ yr}^{-1} \text{ kpc}^{-2}$, comparable to vigorous starburst galaxies with similar SFRs¹⁴, but an order of magnitude or more lower per unit surface area than the extreme starbursts found in high-redshift submillimetre surveys or quasars²⁰.

The comparison of the gas mass to stellar mass in this galaxy shows that about $55 \pm 25\%$ of the baryonic matter is in the form of gas, indicating that the galaxy has already formed much of its stars and metals. Taken together, these lines of evidence point to a picture of A1689-zD1 consistently forming stars at a moderate rate since $z \approx 9$, or possibly having passed through its extreme starburst very rapidly and now being in a declining phase of star formation.

It has been suggested that the decreasing metal contents of high-redshift galaxies will make them challenging to detect at far-infrared wavelengths^{6,7}. A1689-zD1 is at $z = 7.5$ and although it is magnified by a factor of 9.3, it was detected in only brief observations with ALMA. This promises a reasonable detection rate for L^* galaxies in unlensed fields at these redshifts for the full ALMA array, in contrast to the gloomy outlook painted by observations of very-low-metallicity systems^{7,8}. The precise identification and detailed characterization of the star-forming population of the early Universe should therefore be possible in the far-infrared in the near future and should not be restricted to rare hyper-luminous infrared galaxies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 May; accepted 19 December 2014.

Published online 2 March 2015.

1. Smit, R. *et al.* The star formation rate function for redshift $z \sim 4$ –7 galaxies: evidence for a uniform buildup of star-forming galaxies during the first 3 Gyr of cosmic time. *Astrophys. J.* **756**, 14 (2012).
2. Finkelstein, S. L. *et al.* A galaxy rapidly forming stars 700 million years after the Big Bang at redshift 7.51. *Nature* **502**, 524–527 (2013).
3. Shibuya, T. *et al.* The first systematic survey for Ly α emitters at $z = 7.3$ with red-sensitive Subaru/Suprime-Cam. *Astrophys. J.* **752**, 114 (2012).
4. Treu, T. *et al.* The changing Ly α optical depth in the range $6 < z < 9$ from the MOSFIRE spectroscopy of Y-dropouts. *Astrophys. J.* **775**, L29 (2013).
5. Magdis, G. E. *et al.* The molecular gas content of $z = 3$ Lyman break galaxies: evidence of a non-evolving gas fraction in main-sequence galaxies at $z > 2$. *Astrophys. J.* **758**, L9 (2012).
6. Tan, Q. *et al.* A deep search for molecular gas in two massive Lyman break galaxies at $z = 3$ and 4: vanishing CO-emission due to low metallicity? *Astrophys. J.* **776**, L24 (2013).
7. Fisher, D. B. *et al.* The rarity of dust in metal-poor galaxies. *Nature* **505**, 186–189 (2014).
8. Ouchi, M. *et al.* An intensely star-forming galaxy at $z \sim 7$ with low dust and metal content revealed by deep ALMA and HST observations. *Astrophys. J.* **778**, 102 (2013).
9. Bouwens, R. J. *et al.* UV luminosity functions at redshifts $z \sim 4$ to $z \sim 10$: 11000 galaxies from HST legacy fields. *Astrophys. J.* (in the press); preproof at <http://arxiv.org/abs/1403.4295> (2014).
10. Bradley, L. D. *et al.* Discovery of a very bright strongly lensed galaxy candidate at $z \sim 7.6$. *Astrophys. J.* **678**, 647–654 (2008).
11. Chabrier, G. Galactic stellar and substellar initial mass function. *Publ. Astron. Soc. Pacif.* **115**, 763–795 (2003).
12. da Cunha, E. *et al.* On the effect of the cosmic microwave background in high-redshift sub-millimeter observations. *Astrophys. J.* **766**, 13 (2013).
13. Michałowski, M. J., Watson, D. & Hjorth, J. Rapid dust production in submillimetre galaxies at $z > 4$? *Astrophys. J.* **712**, 942–950 (2010).
14. Kennicutt, R. C. & Evans, N. J. Star formation in the Milky Way and nearby galaxies. *Annu. Rev. Astron. Astrophys.* **50**, 531–608 (2012).
15. Draine, B. *et al.* Dust masses, PAH abundances, and starlight intensities in the SINGS galaxy sample. *Astrophys. J.* **663**, 866–894 (2007).

16. Zafar, T. & Watson, D. The metals-to-dust ratio to very low metallicities using GRB and QSO absorbers; extremely rapid dust formation. *Astron. Astrophys.* **560**, A26 (2013).
17. Christensen, L. *et al.* The low-mass end of the fundamental relation for gravitationally lensed star-forming galaxies at $1 < z < 6$. *Mon. Not. R. Astron. Soc.* **427**, 1953–1972 (2012).
18. Smit, R. *et al.* Evidence for ubiquitous high-equivalent-width nebular emission in $z \sim 7$ galaxies: toward a clean measurement of the specific star-formation rate using a sample of bright, magnified galaxies. *Astrophys. J.* **784**, 58 (2014).
19. Stark, D. P. *et al.* Ultraviolet emission lines in young low mass galaxies at $z \sim 2$: physical properties and implications for studies at $z > 7$. *Mon. Not. R. Astron. Soc.* **445**, 3200–3220 (2014).
20. Carilli, C. L. & Walter, F. Cool gas in high-redshift galaxies. *Annu. Rev. Astron. Astrophys.* **51**, 105–161 (2013).
21. Cooray, A. *et al.* HerMES: the rest-frame UV emission and a lensing model for the $z = 6.34$ luminous dusty starburst galaxy HFLS3. *Astrophys. J.* **790**, 40 (2014).
22. Hu, E. M. *et al.* A redshift $z = 6.56$ galaxy behind the cluster Abell 370. *Astrophys. J.* **568**, L75–L79 (2002).
23. Bradley, L. D. *et al.* Through the looking glass: bright, highly magnified galaxy candidates at $z \sim 7$ behind A1703. *Astrophys. J.* **747**, 3 (2012).
24. Iye, M. *et al.* A galaxy at a redshift $z = 6.96$. *Nature* **443**, 186–188 (2006).
25. Tanvir, N. *et al.* A γ -ray burst at a redshift of $z \sim 8.2$. *Nature* **461**, 1254–1257 (2009).
26. Schaerer, D. *et al.* New constraints on dust emission and UV attenuation of $z = 6.5$ – 7.5 galaxies from IRAM and ALMA observations. *Astron. Astrophys.* **574**, A19 (2014).
27. Jiang, L. *et al.* Physical properties of spectroscopically confirmed galaxies at $z \geq 6$. I. Basic characteristics of the rest-frame UV continuum and Ly α emission. *Astrophys. J.* **772**, 99 (2013).
28. Ota, K. *et al.* ALMA observation of 158 μm [C II] line and dust continuum of a $z = 7$ normally star-forming galaxy in the epoch of reionization. *Astrophys. J.* **792**, 34 (2014).
29. Berger, E. *et al.* ALMA observations of the host galaxy of GRB090423 at $z = 8.23$: deep limits on obscured star formation 630 million years after the Big Bang. *Astrophys. J.* (submitted); preprint at <http://arxiv.org/abs/1408.2520> (2014).
30. Tanvir, N. R. *et al.* Star formation in the early universe: beyond the tip of the iceberg. *Astrophys. J.* **754**, 46 (2012).

Acknowledgements The Dark Cosmology Centre is funded by the Danish National Research Foundation. L.C. is supported by the EU under a Marie Curie Intra-European Fellowship, contract number PIEF-GA-2010-274117. K.K. acknowledges support from the Swedish Research Council and the Knut and Alice Wallenberg Foundation. J.R. acknowledges support from a European Research Council starting grant, CALENDs, and the Career Integration Grant 294074. A.G. acknowledges support from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement number 267251 (“AstroFit”). M.J.M. acknowledges the support of the Science and Technology Facilities Council. ALMA is a partnership of the European Southern Observatory (ESO, representing its member states), the National Science Foundation (USA) and National Institutes of Natural Sciences (Japan), together with the National Research Council (Canada) and the National Science Council and the Academia Sinica Institute for Astronomy and Astrophysics (Taiwan), in cooperation with Chile. The Joint ALMA Observatory is operated by the ESO, Associated Universities Inc./National Radio Astronomy Observatory and the National Astronomical Observatory of Japan. We thank L. Lindroos, J. Hjorth, J. Fynbo, A. C. Andersen, and R. Bouwens for discussions, M. Limousin for providing a lensing map of the cluster, and the Nordic ALMA Regional Center Node for assistance.

Author Contributions D.W. conceived the study, was Principal Investigator of the X-shooter programme, produced Fig. 1 and Extended Data Figs 1 and 4–7 and wrote the main text. L.C. reduced and analysed the X-shooter spectrum, did the HyperZ analysis and produced Fig. 2 and Extended Data Fig. 2. K.K. reduced and analysed the ALMA data and produced Fig. 3 and Extended Data Fig. 3. J.R. was Principal Investigator of the ALMA programmes and reduced and analysed the Hubble data. A.G. modelled the ultraviolet SED and determined the galaxy stellar age. M.J.M. modelled the full ultraviolet–far-infrared SED and produced Table 1. All authors contributed to the Methods and all authors discussed the results and commented on the manuscript.

Author Information This paper makes use of the following ALMA data: ADS/JAO.ALMA 2011.0.00319.S and 2012.1.00261.S available from the ALMA archive at <https://almascience.eso.org/alma-data/archive>. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.W. (darach@dark-cosmology.dk).

METHODS

All uncertainties quoted are 68% confidence unless stated otherwise.

Optical spectroscopy. The X-shooter spectrograph on the Very Large Telescope was used to observe the source A1689-zD1 on the nights 21 March and 18 April 2010, as well as shorter exposures on 27 March, 25 April, 1, 2 and 3 May 2011, and 17 and 24 March 2012. The observations used a standard dither pattern. Approximately half the exposure time (the 2010 observations) was performed with a fixed slit position angle, while the rest used the parallactic angle. The slit setups are indicated in Fig. 1. The data were reduced with the ESO X-shooter pipeline version 2.2. The standard calibration routines were used, with the different nod positions employed to subtract the background sky emission. Spectrophotometric standard stars observed on the same night as the galaxy were used for flux calibration. The target was acquired using an offset from a bright nearby star, calculated from the Hubble Space Telescope imaging. Offsets are executed with an accuracy of $<0.1''$, that is, the location of the galaxy in the slit is known a priori to better than a pixel. The slit-centre position and offsets along the slit were used to shift and co-add the data into a final combined two-dimensional spectrum.

Spectroscopic redshift. The spectroscopy rules out a low-redshift solution. We detect no emission lines in the spectrum (see below), but the stellar continuum of the galaxy is detected, allowing us to measure the blue end of the ultraviolet continuum and a sharp break, which cannot be reproduced with dust extinction. Other breaks are excluded by the sharpness and depth of the break in spectroscopy, the blue slope of the ultraviolet continuum, and the deep upper limits in the VIS arm of the spectrograph. These data are independent of the original discovery data for the dropout, removing any Eddington-bias-like effect and improving the reliability of the redshift determination.

The SNR per spectral pixel is low. We therefore bin the X-shooter spectra by a large factor in wavelength (see ref. 31 for details). We fit a template spectrum of a ~ 100 -million-year-old galaxy to estimate the redshift. This template was shifted to a range of redshifts and corrected for the transmission in the intergalactic medium³². The best-fit redshift was found from χ^2 minimization. We tested the fit with a variety of bin-sizes: 150, 180, 200, 250 and 400 pixels in the near-infrared. The VIS-spectrum data bin sizes are 1.5 times larger. We extracted the one-dimensional spectrum for each binning and performed 1,000 realizations of the model where we added Gaussian noise corresponding to the uncertainty in each binned pixel. We calculated the average of the five binning sizes and determined the average 68% confidence level to be $z = 7.5 \pm 0.2$.

To be confident of the location of the spectral break, we also adopt a standard methodology for detecting steps in one-dimensional data: searching for a change in slope in the unbinned cumulative sum (Extended Data Fig. 1). The break at $\sim 1,035$ nm is clearly observed only in the data from the NIR arm of the spectrograph. We simulated the data 500 times, using the error spectrum as the standard deviation for the Gaussian random realizations; following the same cumulative sum method and data-cleaning as for the real data, we derived the break position and its uncertainty. This is very close to the analysis of the binned data: $z = 7.5 \pm 0.2$. The fit results and error are independent of the VIS arm data and depend only on the NIR spectrum.

A galaxy at $z \approx 2$ is reported¹⁰ offset by $1.5''$ from the lensed galaxy. In Extended Data Fig. 2 we show its undiluted spectrum from the 2010 data alone. Its spectrum extends across the NIR/VIS arm boundary and does not show the sharp break seen in A1689-zD1. In Fig. 2, the spectrum of this nearby galaxy is diluted because the spectra are observed at different slit orientations, some of which did not cover the companion (see Fig. 1).

Finally, we fit the full set of photometric data (all Hubble and Spitzer bands) with the New-HyperZ code³³, version 12.2, to determine a photometric redshift, for completeness. Only one redshift solution is allowed, at $z \approx 7.4$ (Extended Data Fig. 3), consistent with the spectroscopic determination.

Emission line flux detection limits. No emission line is found from visual inspection of the spectrum. Of the strong emission lines, at $z = 7.5$ only Ly α could in principle be detected (the He II 1,640 Å line would be too faint). We estimated the limits for the detection of any emission line by adding artificial emission lines at random within the redshift range determined from the spectral break redshift calculation. These lines were simulated with FWHMs in the range $50\text{--}200\text{ km s}^{-1}$ and were added to the unbinned two-dimensional spectrum. The one-dimensional spectra were extracted and binned by 3–5 pixels. A Gaussian function was fitted to the data and the SNR of the artificial line flux was derived.

We place a 3σ upper limit to the Ly α emission line of $1.8 \times 10^{-17}\text{ erg cm}^{-2}\text{ s}^{-1}$ at $10,030\text{ Å}$ near the sky lines and a factor of seven lower away from the sky lines (not corrected for lensing magnification). This limit becomes tighter from the blue to the red, that is, at $10,330\text{ Å}$ and $10,573\text{ Å}$ the limits near the sky lines are $0.9 \times 10^{-17}\text{ erg cm}^{-2}\text{ s}^{-1}$ and $0.5 \times 10^{-17}\text{ erg cm}^{-2}\text{ s}^{-1}$, respectively.

To determine the effective Ly α escape fraction, we first convert this limit to a nominal SFR by assuming a factor-of-8.7 flux ratio between Ly α and H α , and a

standard conversion between H α and the SFR³⁴, but assuming a Chabrier initial mass function¹¹, as used throughout the paper. This corresponds to a detection limit for the nominal Ly α SFR $<0.7\text{ M}_{\odot}\text{ yr}^{-1}$ (corrected for lensing). Between the sky lines, the upper limit is $<0.09\text{ M}_{\odot}\text{ yr}^{-1}$. This implies an effective Ly α escape fraction of either $<6\%$ or $<0.8\%$ by comparison to the total SFR of $12\text{ M}_{\odot}\text{ yr}^{-1}$. The rest-frame Ly α equivalent width is $<27\text{ Å}$, or $<4\text{ Å}$ between the sky lines. We cannot distinguish between Ly α absorbed by the galaxy interstellar medium, circumgalactic medium or intergalactic medium. However, the dustiness of the galaxy indicates that a substantial fraction of the Ly α may be absorbed in the host.

Similarly, we place a 3σ upper limit of $2 \times 10^{-18}\text{ erg cm}^{-2}\text{ s}^{-1}$ for the lensing uncorrected flux of the C III] 1,909 Å emission line, which lies in a region of the spectrum relatively free of atmospheric effects, corresponding to a restframe equivalent width of $<4\text{ Å}$.

The ALMA observations and data reduction. We obtained ALMA mosaic observations for A1689 in Cycle 0 and Cycle 1 as part of the projects 2011.0.00319.S and 2012.1.00261.S. The receivers were tuned to cover 211.06–214.94 GHz, 221.46–225.34 GHz, 227.06–230.94 GHz and 237.46–241.34 GHz; the 2012.1.00261.S data used here cover the first and third frequency setup. The correlator was used in the frequency domain mode with a bandwidth of 1,875 MHz in each spectral window. The projected baselines range between 12 m and 450 m. The quasar 3C 279 was used for bandpass and phase calibration. The distance from 3C 279 to A1689 is 5.9° . Mosaicked images were primary-beam-corrected and the weights of the individual fields were taken into account. Flux calibration was done using Mars and Titan.

Data reduction used the Common Astronomy Software Application³⁵ (<http://casa.nrao.edu/>) versions 3.4 and 4.1 for the Cycle 0 and 1 data respectively. For the Cycle 0 data, additional careful reduction of the observatory-provided preliminary-reduced data was required, including flagging of noisy data as well as producing a frequency-dependent model of 3C 279 to account for the spectral index of the continuum emission through bootstrapping from the flux calibrators. Our results are, however, not more than 1σ different from the preliminary-reduced data.

The data were combined and imaged using CASA 4.2; when imaging, the calibrated visibilities were naturally weighted (resulting beam size $0.8'' \times 1''$) and tapered using a two-dimensional Gaussian with $1'' \times 1''$ to give greater weight to shorter baselines, resulting in a beam-size of $1.36'' \times 1.15''$. The full mosaic will be presented in a forthcoming paper (K.K. *et al.*, manuscript in preparation). The SNR images are shown in Fig. 3. The flux image for A1689-zD1 is shown in Extended Data Fig. 4.

The most conservative estimate of the astrometric uncertainty is half the beam dimensions, that is, $0.35'' \times 0.5''$, dominated by residual atmospheric phase effects and the SNR. Together with four other ALMA-detected sources, we found an average offset of $0.4\text{--}0.45''$ between the ALMA centroids and the Hubble centroids. Part of these offsets is probably contributed by intrinsically different optical and far-infrared morphologies. No noticeable flux is detected from the galaxy $1.5''$ from A1689-zD1.

No spectral line is detected towards A1689-zD1 in the ALMA data. Our data cover [C II] 158 μm emission at: $z = 6.87\text{--}7.00$, $7.23\text{--}7.37$, $7.43\text{--}7.58$ and $7.84\text{--}8.00$. This covers $\sim 50\%$ of the 1σ and 2σ ranges allowed by X-shooter. The typical observed luminosity ratio $L_{[\text{C II}]} / L_{\text{FIR}}$ varies between 0.001 and 0.008 (ref. 36). Over the redshift ranges above, we exclude [C II] 158 μm emission at the high end of this ratio (>0.0024 , 5σ for a linewidth of 100 km s^{-1}).

Code availability. The data reduction scripts used for the ALMA data are freely available and can be obtained at <http://dark.nbi.ku.dk/research/archive>.

Dust mass and temperature. We derive the far-infrared luminosity (42–122 μm restframe) and dust mass using a single temperature modified blackbody and assuming a typical value for the dust mass absorption coefficient of $\kappa_{\nu} = 0.067 \times \left[\frac{\nu}{250} \right]^{\beta_{\text{IR}}}\text{ m}^2\text{ kg}^{-1}$ at the observed frequency $\nu = 226\text{ GHz}$, and $\beta_{\text{IR}} = 1.92$ (as observed in the high-redshift galaxy HFLS3³⁷). Different values for the dust mass absorption coefficient could result in dust masses $\pm 0.18\text{ dex}$ (ref. 38). We explore the effect of using more complex full SED models in a section below.

With only one far-infrared flux point it was necessary to assume a value of the dust temperature to recover the dust mass. We assumed 35 K. A lower temperature would increase the inferred dust mass, making the galaxy richer in dust than we infer. The galaxy may not be optically thin to far-infrared radiation; we assume the wavelength of optical depth unity is $\lambda_0 = 200\text{ }\mu\text{m}$. We apply corrections due to CMB heating and the error induced by the large CMB background subtraction¹². We show the SFR and dust mass inferred using our fiducial modified blackbody model for various assumed dust temperatures, β_{IR} and λ_0 values in Extended Data Fig. 5.

The constraints induced by the observed correlation between the ultraviolet spectral slope, β_{UV} , and the ratio of observed infrared to ultraviolet luminosity, IRX^{39} , with the flux uncertainties and β_{UV} slope uncertainties included, are also shown in Extended Data Fig. 5. The dust mass, M_{D} , for the fiducial model is $\log(M_{\text{D}}/M_{\odot}) = 7.6$, with a 2σ lower bound of 7.30, and an upper bound that is not

strongly constrained, but where $\log(M_D/M_\odot) = 7.9$ implies a dust-to-gas ratio approximately five times that of the Milky Way (see below), and is therefore disfavoured. It has been shown recently that galaxies at high redshift may be preferentially more dusty⁴⁰. This higher IRX, by 0.3 dex, is quite consistent with A1689-zD1.

Star-formation rate. We measure the ultraviolet flux F_ν from the Hubble Space Telescope F160W photometric data point. The ultraviolet luminosity is calculated for the best-fit redshift of 7.5 using our standard cosmological parameters⁴¹. No dust correction is applied in calculating the ultraviolet SFR. The SFR is derived from the ultraviolet and total infrared (3–1,100 μm) luminosities³³, L_{UV} and L_{TIR} . For the ultraviolet luminosity we find $L_{UV} = 1.8 \pm 0.2 \times 10^{10} L_\odot$, corresponding to a SFR $= 2.7 \pm 0.3 M_\odot \text{ yr}^{-1}$. For the total infrared luminosity we obtain $L_{TIR} = 6.2 \pm 0.8 \times 10^{10} L_\odot$, which is the flux uncertainty only, using our fiducial model, but where the true uncertainty is constrained primarily by the desire to remain consistent within the scatter with the IRX– β_{UV} relation. These values are consistent with those derived from modelling the SED (see below) and result in $\text{SFR}_{TIR} = 9 \pm 2 M_\odot \text{ yr}^{-1}$, using recent calibrations based on the 3–1,100 μm infrared luminosity¹⁴. A SFR_{TIR} much below $7 M_\odot \text{ yr}^{-1}$ is excluded in this model. A substantially lower dust temperature could result in a higher SFR_{TIR}, but would imply an excessive dust mass. From the analysis in Extended Data Fig. 5, our best estimate is $\text{SFR}_{TIR} = 9^{+4}_{-2} M_\odot \text{ yr}^{-1}$.

The sum of the ultraviolet and infrared SFRs results in a total SFR of $12^{+4}_{-2} M_\odot \text{ yr}^{-1}$, for this initial mass function (statistical uncertainty only), with a likely additional contribution to the luminosity and hence the SFR from the mid-infrared emission that is not well-modelled by a single temperature modified blackbody (see SED modelling below). Therefore, we take this as a conservative estimate of the SFR and note that it may be somewhat higher.

Gas mass and metallicity. We invert the Schmidt–Kennicutt law relating the surface densities of SFR and gas mass¹⁴ to obtain constraints on the gas mass. The dominant uncertainty on this conversion comes from the uncertainty in the slope of the Schmidt–Kennicutt law, which is a ± 0.4 dex scatter—the dependence on the surface area of the galaxy is very small (0.04 dex) since it appears in both the SFR and gas mass terms. We here use the relation as originally derived, which reproduces even starburst galaxies quite well, assuming a constant carbon monoxide (CO) to gas conversion, $X(\text{CO})$ ¹⁴. Low-metallicity galaxies may have a lower gas mass for a given SFR¹⁴, which would only increase the dust-to-gas ratio derived here. The derived gas mass is inversely dependent on the inferred lensing magnification. Since A1689-zD1 is not very close to the critical lines, its lensing amplification is likely to be fairly accurate. A different analysis of the data for the cluster gives a similar magnification of 8.6 ± 0.5 for this source (M. Limousin, private communication), confirming that the lens model contributes only a small additional uncertainty, 0.025 dex, for the calculations presented here. The overall uncertainty in the gas mass summed in quadrature is therefore 0.45 dex, including the scatter in the Schmidt–Kennicutt relation and the uncertainty on the SFR.

The dust-to-gas ratio we derive for our fiducial model is 0.017, with a total uncertainty of 0.5 dex under the assumptions outlined above for the derivation of the dust and gas masses. Both values depend on the SFR and this is accounted for in the uncertainty on the ratio. The scatter in the Schmidt–Kennicutt relation and the dust mass systematics dominate the uncertainty.

The age, mass and SED of A1689-zD1. We construct the photometric SED of A1689-zD1 from the Hubble Space Telescope’s ACS camera F775W and F850W, WFC3 camera F105W, F125W, F140W and F160W, and NIC2 camera F110W bands, and the Spitzer Space Telescope’s IRAC camera 3.6 μm and 4.5 μm bands. Since its first discovery with the Hubble Space Telescope, A1689-zD1 was observed on April 2010 with the Hubble WFC3 in the F105W, F125W, F140W and F160W bands, for 2.5 ks each (proposal ID 11802, Principal Investigator: H. Ford). The individual reduced frames were combined using the MultiDrizzle⁴² software onto a common pixel scale of $0.05''$ per pixel and the astrometry was matched to the ACS images. Isophotal magnitudes have been measured using SExtractor in double-image mode, taking the F140W image as a reference for the detection. The fluxes for the source from the newer Hubble imaging are: for F105W, 25.49 ± 0.18 , for F125W, 25.00 ± 0.13 , for F140W, 24.64 ± 0.05 , and for F160W, 24.51 ± 0.11 . In addition, photometry from ref. 10 is used.

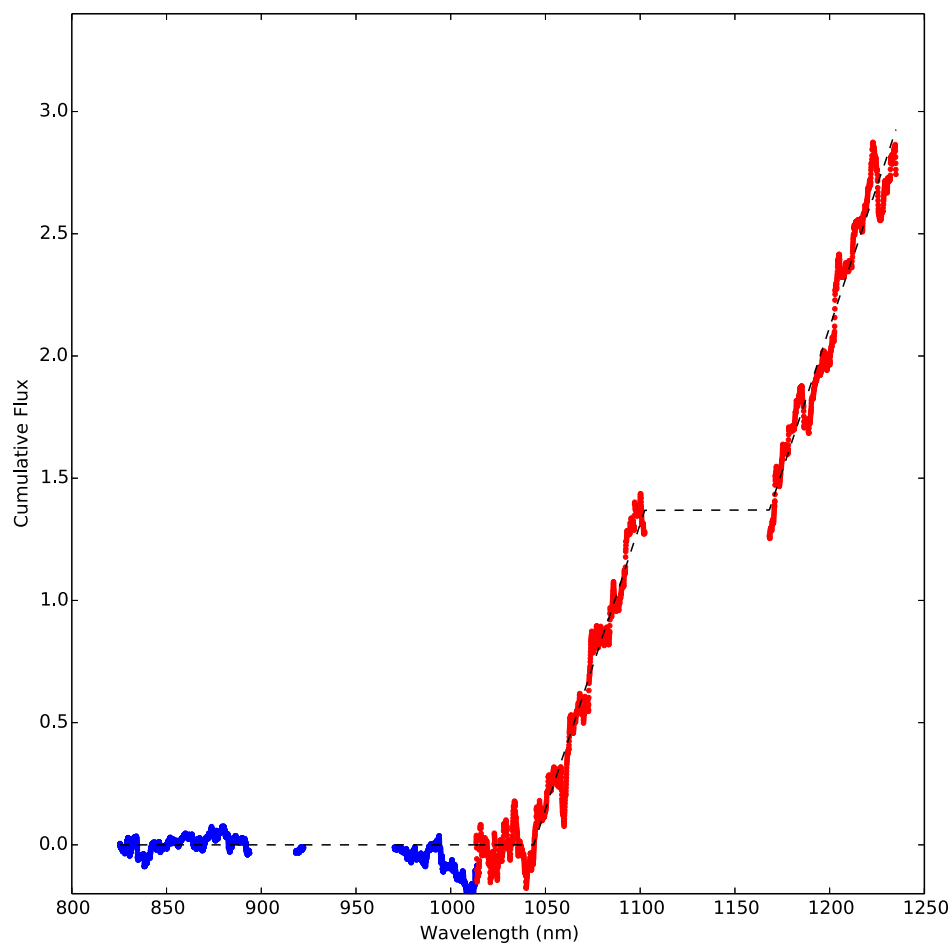
We build a large library of model galaxy SEDs by convolving Simple Stellar Population models⁴³ with randomly generated star-formation histories. These are modelled with exponentially declining laws (parameterized by the formation redshift—bound to be younger than the Universe’s age at $z = 7.5$ —and the star-formation timescale) with stochastic bursts of star formation (parameterized by their duration and fraction of mass formed). The models span a metallicity range between 20% and 2.5 times solar. Dust attenuation is applied to the stellar SEDs adopting a two-component model⁴⁴ and randomly generating the values for the total optical depth and the fraction of it contributed by the diffuse interstellar medium. Finally, transmission in the intergalactic medium is applied³².

Following the Bayesian approach developed and extensively applied to lower-redshift data^{45,46}, the observed SED is compared to every model SED in the library. This allows us to construct the full probability density function of parameters of interest by weighting each model by its likelihood $\exp(-\chi^2/2)$ and marginalizing over the nuisance parameters of the star-formation histories. For completeness we also derive the total (dust-corrected and averaged over the galaxy age) SFR, $9^{+5}_{-3} M_\odot \text{ yr}^{-1}$, and the dust attenuation, $A_{1600} = 1.0 \pm 0.4$ mag, consistent with the values derived elsewhere in the paper. Because of the almost complete absorption of the flux below the Ly α by the intergalactic medium, whether we include the ACS upper limits or not does not change the fit. The best fit to the ultraviolet–optical SED is shown in Extended Data Fig. 6.

Finally, strong emission lines are not included in the SED modelling. We consider possible contamination by the Balmer H γ and H δ lines to the 3.6 μm flux and the contamination by the H β and [O III] lines to the 4.5 μm flux. Adopting a SFR of $12 M_\odot \text{ yr}^{-1}$ and standard case B Balmer line ratios, we expect the 3.6 μm flux contribution to be at most 7%. The nebular line contribution to the 4.5 μm flux could amount to 8%–14% (depending on the [O III]/H β line ratio). By applying these corrections to the observed fluxes, our SED fit would yield a stellar mass 0.06 dex lower and a light-weighted mean age 0.04 dex younger than obtained above.

As a final check, we fit self-consistent full ultraviolet to far-infrared SED models using the MAGPHYS⁴⁷ and GRASIL^{48–50} codes. From these fits we derived the following lensing-corrected parameters: $\text{SFR} = 10^{+4}_{-2} M_\odot \text{ yr}^{-1}$, $\log(M_\star/M_\odot) = 9.3^{+0.2}_{-0.1}$, $M_D = 3^{+3}_{-2} \times 10^7 M_\odot$ (MAGPHYS); $\text{SFR} = 9^{+34}_{-2} M_\odot \text{ yr}^{-1}$, $\log(M_\star/M_\odot) = 9.4^{+0.6}_{-0.2}$, $M_D = 7^{+2}_{-2} \times 10^7 M_\odot$ (GRASIL). Applying the same order of CMB correction to these fits as to the modified blackbody would increase each parameter by about 30% (that is, +0.12 dex). The fits with these codes are shown in Extended Data Fig. 7.

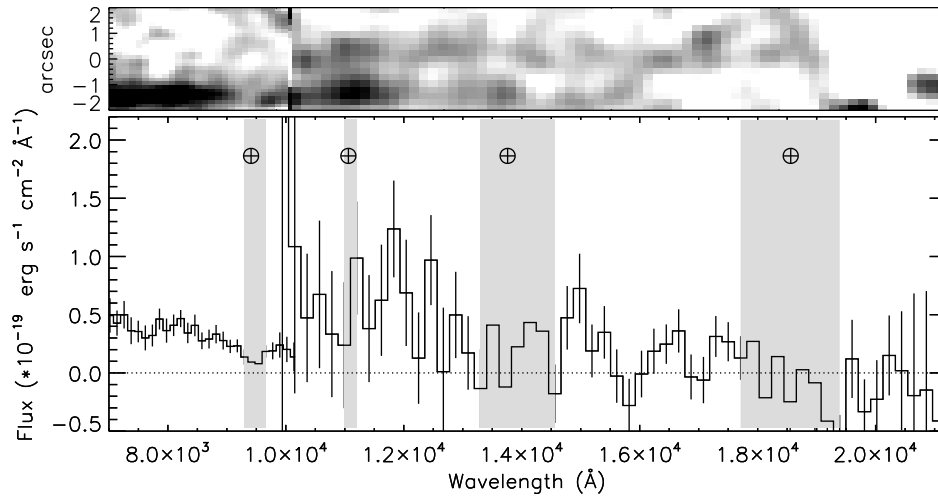
- Christensen, L. *et al.* Gravitationally lensed galaxies at $2 < z < 3.5$: direct abundance measurements of Ly α emitters. *Mon. Not. R. Astron. Soc.* **427**, 1973–1982 (2012).
- Madau, P. Radiative transfer in a clumpy universe: the colors of high-redshift galaxies. *Astrophys. J.* **441**, 18–27 (1995).
- Bolzonella, M., Miralles, J.-M. & Pelló, R. Photometric redshifts based on standard SED fitting procedures. *Astron. Astrophys.* **363**, 476–492 (2000).
- Kennicutt, R. C. Star formation in galaxies along the Hubble sequence. *Annu. Rev. Astron. Astrophys.* **36**, 189–231 (1998).
- McMullin, J. P., Waters, B., Schiebel, D. & Young, W. in *Astronomical Data Analysis Software And Systems XVI* (eds Shaw, R. A., Hill, F. & Bell, D. J.) *Astron. Soc. Pacif. Conf. Ser.* **376**, 127–130 (2007).
- Díaz-Santos, T. Explaining the [C II] 157.7 μm deficit in luminous infrared galaxies—first results from a *Herschel*/PACS study of the GOALS sample. *Astrophys. J.* **774**, 68 (2013).
- Riechers, D. A. *et al.* A dust-obscured massive maximum-starburst galaxy at a redshift of 6.34. *Nature* **496**, 329–333 (2013).
- Andersen, A. C. in *Why Galaxies Care About AGB Stars: their Importance as Actors And Probes* (eds Kirshenbaum, F., Charbonnel, C. & Wing, R. F.) *ASP Conf. Ser.* **768**, 170–180 (2007).
- Meurer, G. R., Heckman, T. M. & Calzetti, D. Dust absorption and the ultraviolet luminosity density at $z \sim 3$ as calibrated by local starburst galaxies. *Astrophys. J.* **521**, 64–80 (1999).
- Oteo, I. Dust correction factors over $0 < z < 3$ in massive star-forming galaxies from a stacking analysis of *Herschel* data. *Astron. Astrophys.* **572**, L4 (2014).
- Komatsu, E. *et al.* Seven-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological interpretation. *Astrophys. J.* **192** (Suppl.), 18 (2011).
- Koekemoer, A. M., Fruchter, A. S., Hook, R. N. & Hack, W. in *The 2002 HST Calibration Workshop: HST after the Installation of the ACS and the NICMOS Cooling System* (eds Arribas, S., Koekemoer, A. & Whitmore, B.) 337–340 (Space Telescope Science Institute, 2003); http://www.stsci.edu/hst/HST_overview/documents/calworkshop/workshop2002/CW2002_Papers/koekemoer_multidrizze.pdf.
- Bruzual, G. & Charlot, S. Stellar population synthesis at the resolution of 2003. *Mon. Not. R. Astron. Soc.* **344**, 1000–1028 (2003).
- Charlot, S. & Fall, M. A simple model for the absorption of starlight by dust in galaxies. *Astrophys. J.* **539**, 718–731 (2000).
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M. & Tremonti, C. A. The ages and metallicities of galaxies in the local universe. *Mon. Not. R. Astron. Soc.* **362**, 41–58 (2005).
- Salim, S. *et al.* New constraints on the star formation histories and dust attenuation of galaxies in the local Universe from GALEX. *Astrophys. J.* **619**, L39–L42 (2005).
- da Cunha, E., Charlot, S. & Elbaz, D. A simple model to interpret the ultraviolet, optical and infrared emission from galaxies. *Mon. Not. R. Astron. Soc.* **388**, 1595–1617 (2008).
- Silva, L., Granato, G. L., Bressan, A. & Danese, L. Modeling the effects of dust on galactic spectral energy distributions from the ultraviolet to the millimeter band. *Astrophys. J.* **509**, 103–117 (1998).
- Iglesias-Páramo, J. *et al.* UV to IR SEDs of UV-selected galaxies in the ELAIS fields: evolution of dust attenuation and star formation activity from $z = 0.7$ to 0.2. *Astrophys. J.* **670**, 279–294 (2007).
- Michalowski, M., Hjorth, J. & Watson, D. Cosmic evolution of submillimeter galaxies and their contribution to stellar mass assembly. *Astron. Astrophys.* **514**, A67 (2010).



Extended Data Figure 1 | Cumulative sum of the unbinned spectrum.

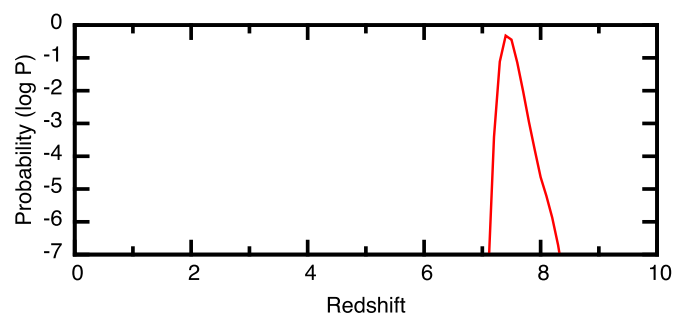
The VIS and NIR arms are plotted in blue and red respectively. The best-fitting step function is plotted as a dashed line. The break in the spectrum is clearly

detected with the NIR arm only. Gaps in the cumulative spectrum are due to removal of regions affected by strong sky absorption.

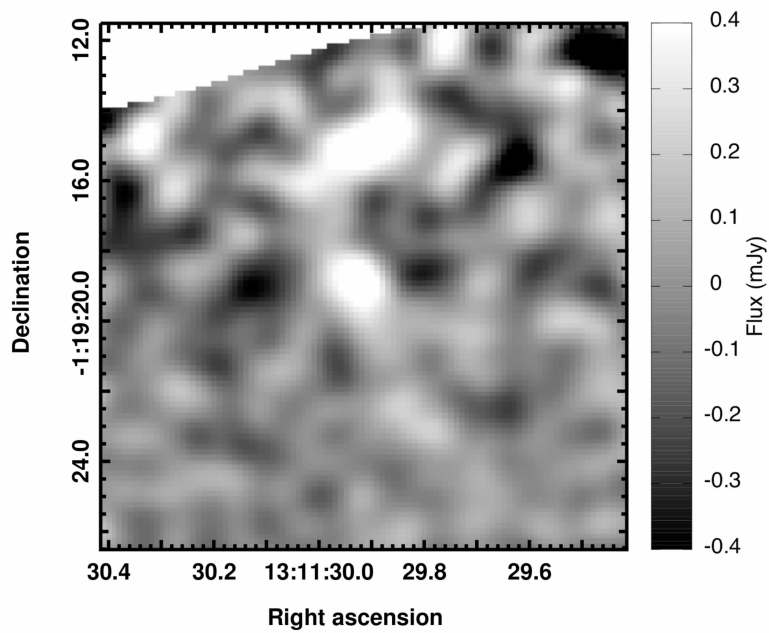


Extended Data Figure 2 | Spectrum obtained only at position angle 64° East of North. The slit consistently covered both the emission from the high redshift galaxy and the galaxy located $2''$ below it. This spectrum uses approximately half of the total exposure time. The upper panel shows the two-dimensional rectified spectrum, the lower panel the one-dimensional

spectrum of the companion. Error bars are 68% confidence. The spectrum of the companion galaxy is recovered through the entire spectral range, including that covered by the transition from the VIS to the NIR data, and shows no indication of the sharp break seen in A1689-zD1.

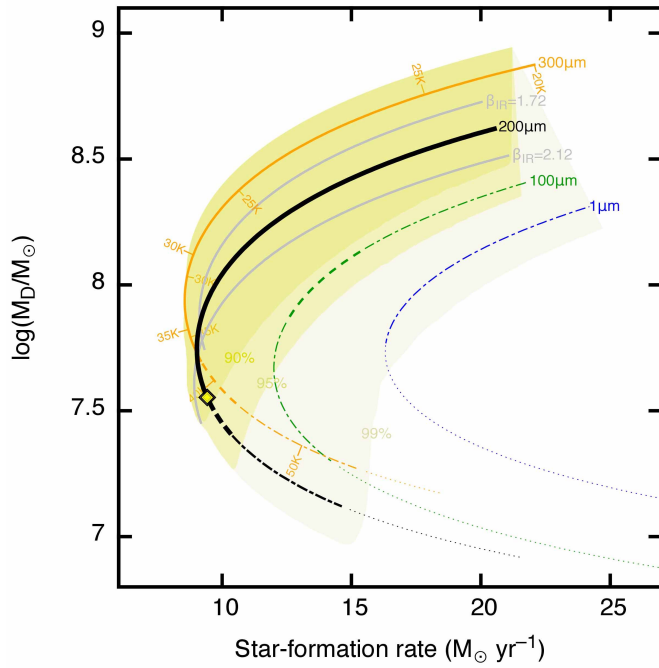


Extended Data Figure 3 | Probability distribution as a function of redshift for galaxy template fits to the Hubble and Spitzer IRAC photometry data. The probability distribution is based on fitting galaxies using the New-HyperZ code³³.

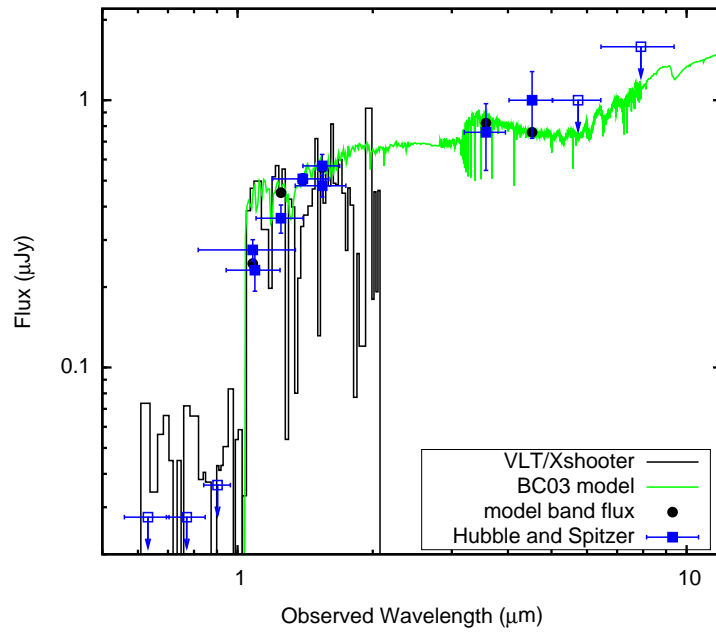


Extended Data Figure 4 | The tapered ALMA flux image at 226 GHz, centred on A1689-zD1; the image is primary-beam-corrected. The depth of the map at the location of A1689-zD1 is 0.12 mJy per beam (42% of the deepest part of the mosaic). The sensitivity decreases towards the edge of the mosaic

owing to the overlap of multiple pointings and primary beam correction. The structure north of A1689-zD1 is a probable detection of a different source in the field and will be presented in a forthcoming paper (K.K. *et al.*, manuscript in preparation).

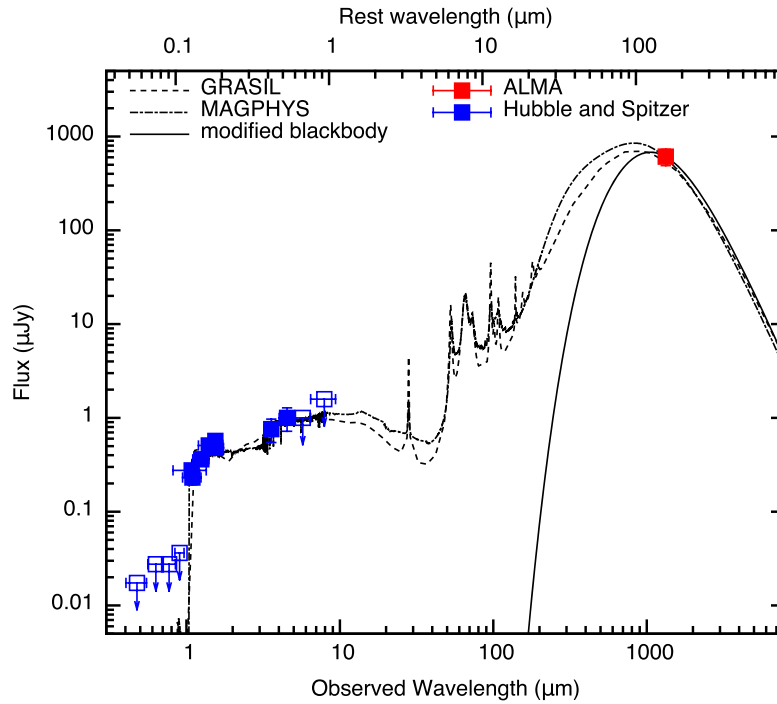


Extended Data Figure 5 | Dust mass and SFR_{TIR} from modified blackbody fits. Tracks show how the parameters change with temperature, with different tracks for different opacity wavelengths, λ_0 . Varying β_{IR} is shown for $\lambda_0 = 200 \mu\text{m}$ (black and grey lines). Intrinsic (CMB-corrected) and measured temperatures are indicated for $\lambda_0 = 300 \mu\text{m}$ (orange line) on the concave and convex sides respectively. A diamond marks our fiducial model: uncorrected $T = 35 \text{ K}$, $\lambda_0 = 200 \mu\text{m}$, $\beta_{\text{IR}} = 1.92$. Solid-colour regions show $<90\%$, $<95\%$, and $<99\%$ confidence intervals due to the $\beta_{\text{UV}}\text{--IRX}$ relation (including measurement uncertainties, $\beta_{\text{IR}} = 1.72\text{--}2.12$, and $\lambda_0 < 300 \mu\text{m}$), with solid, dashed, and dot-dashed lines indicating these intervals for the tracks. Dotted lines mark $>99\%$.



Extended Data Figure 6 | Ultraviolet-optical SED for A1689-zD1. Stellar synthesis models from ref. 43 (BC03) are fitted to the photometric data (squares). Error bars are 68% confidence. The best-fitting model is shown in

green with the resultant fluxes in the different bands shown as circles. The Very Large Telescope (VLT)/X-shooter spectrum is also plotted (solid histogram) for comparison.



Extended Data Figure 7 | SED of A1689-zD1. Full, self-consistent ultraviolet-to-far-infrared models are fitted to the data using the GRASIL (dashed line) and MAGPHYS (dash-dotted line) codes. The values derived from these models fitted to the photometric data (squares) are largely consistent with those derived from the modified blackbody (solid line) and

ultraviolet-optical-only fit, though with an additional contribution from the restframe mid-infrared flux. A CMB correction has not been applied here. Error bars are 68% confidence. Upper limits are 68% confidence for all points except the 8.0- μm band, for which the upper limit is 95% confidence.

Highly efficient star formation in NGC 5253 possibly from stream-fed accretion

J. L. Turner¹, S. C. Beck², D. J. Benford³, S. M. Consiglio¹, P. T. P. Ho⁴, A. Kovács⁵, D. S. Meier^{6,7} & J.-H. Zhao⁸

Gas clouds in present-day galaxies are inefficient at forming stars. Low star-formation efficiency is a critical parameter in galaxy evolution: it is why stars are still forming nearly 14 billion years after the Big Bang¹ and why star clusters generally do not survive their births, instead dispersing to form galactic disks or bulges². Yet the existence of ancient massive bound star clusters (globular clusters) in the Milky Way suggests that efficiencies were higher when they formed ten billion years ago. A local dwarf galaxy, NGC 5253, has a young star cluster that provides an example of highly efficient star formation³. Here we report the detection of the $J = 3 \rightarrow 2$ rotational transition of CO at the location of the massive cluster. The gas cloud is hot, dense, quiescent and extremely dusty. Its gas-to-dust ratio is lower than the Galactic value, which we attribute to dust enrichment by the embedded star cluster. Its star-formation efficiency exceeds 50 per cent, tenfold that of clouds in the Milky Way. We suggest that high efficiency results from the force-feeding of star formation by a streamer of gas falling into the galaxy.

The Submillimeter Array image of NGC 5253, shown in Fig. 1, reveals a bright CO($3 \rightarrow 2$) source coincident with the giant cluster and its ‘supernova’⁴. ‘Cloud D’ (ref. 3) is one of only two molecular clouds detected within the galaxy; the second cloud is smaller and located $\sim 5''$ (90 pc) to the southwest. A ‘streamer’ of gas extending along the minor axis is also detected in CO($3 \rightarrow 2$). This streamer, previously detected in lower- J CO lines, seems to be falling into the galaxy near the supernova^{3,5}. Both the streamer and Cloud D emit 870- μm continuum emission, as shown in Fig. 2. Also shown is an image of 350- μm continuum, in which both Cloud D and the streamer are detected.

The molecular gas in Cloud D is hot. This is clear from the increase in brightness from CO($2 \rightarrow 1$) (ref. 3) to CO($3 \rightarrow 2$). The intensity ratio of the two lines is $I_{32}/I_{21} = 2.6 \pm 0.5 (I_{\text{line}} = \int T_{\text{line}} dv)$. This ratio is non-thermal, although the thermal limit of 2.25 ($\sim v^2$) is within the uncertainties and is what we adopt. Non-local-thermodynamic-equilibrium (non-LTE) modelling of this ratio using RADEX⁶ indicates a minimum kinetic temperature of $T_K > 200$ K for the 1σ lower limit, and $T_K > 350$ K for the adopted value of $I_{32}/I_{21} = 2.25$ (see Methods). The high gas temperature is consistent with a thermal origin for H₂ 2.2- μm emission in the region⁷. Cloud D seems to be a photon-dominated region, heated by ultraviolet radiation from the several thousand cluster O stars in the cluster^{4,8}. The CO($3 \rightarrow 2$)-emitting gas is dense, with $n_{\text{H}_2} \approx (4.5 \pm 0.5) \times 10^4 \text{ cm}^{-3}$.

By contrast with Cloud D, the streamer consists of more typical cool giant molecular clouds. Its value of $I_{32}/I_{21} = 1.0 \pm 0.3$ is consistent with optically thick emission, for which RADEX models allow temperatures as low as $T_K \approx 15\text{--}20$ K, and number densities $n_{\text{H}_2} \approx 3.5\text{--}4 \times 10^4 \text{ cm}^{-3}$. The mass of the streamer is $M_{\text{H}_2} \approx 2 \times 10^6 M_\odot$ ³, which is 1–2% of the stellar mass of the galaxy. The streamer is molecular, dense, and primed for star formation, even before entering the galaxy. This is unlikely to be a primordial collapsing filament⁹, but is more probably previously enriched gas.

The star-formation efficiency of a cloud or region can be defined as $\eta = M_{\text{stars}}/(M_{\text{gas}} + M_{\text{stars}})$, where M_{stars} is the stellar mass and M_{gas} the mass of molecular gas. M_{gas} can be hard to define for star-forming regions within giant molecular clouds, but the association of the isolated Cloud D with the supernova gives us an opportunity to calculate η directly for the giant molecular cloud giving birth to this massive star cluster—if we can determine the mass of Cloud D.

CO is often used to estimate the mass of molecular gas, but this is unreliable in NGC 5253. Here we use the width of the CO line to determine a gas mass for Cloud D based on dynamical considerations. The CO linewidth is $\sigma = 9.2 \pm 0.6 \text{ km s}^{-1}$, based on a Gaussian fit. The cloud dimensions, deconvolved from the beam, are $2.8'' \times 1.5'' \pm 0.1''$ (52 pc \times 28 pc). The virial mass is $M_{\text{vir}} < 1.8^{+0.2}_{-0.7} \times 10^6 M_\odot$ for Cloud D, with uncertainties due to the unknown internal mass distribution (see Methods).

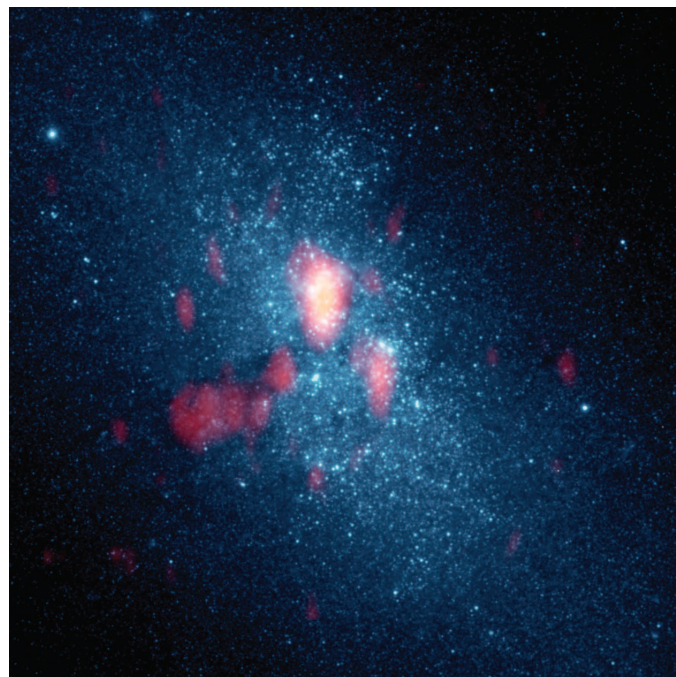


Figure 1 | CO $J = 3 \rightarrow 2$ emission in NGC 5253. The Submillimeter Array (SMA) CO($3 \rightarrow 2$) integrated line intensity, in red, is shown atop a $\lambda 814 \text{ nm}$ Hubble Space Telescope image. The SMA beam is $4'' \times 2''$ (74 pc \times 37 pc). The field covers $40'' \times 40''$ (740 pc \times 740 pc), north up, east left. Image registration is to less than $1''$. The CO streamer coincides with the optical dust lane to the east. The massive star cluster is located at the bright, compact CO peak, Cloud D; it is embedded^{8,9} and is not visible here. Cloud F is to the southwest of Cloud D.

¹Department of Physics and Astronomy, University of California, Los Angeles, Los Angeles, California 90095-1547, USA. ²Department of Physics and Astronomy, University of Tel Aviv, 69978 Ramat Aviv, Israel. ³Observational Cosmology Laboratory, Code 665, NASA at Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. ⁴Academia Sinica, Astronomy and Astrophysics, 11F Astronomy-Mathematics Building, AS/NTU No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan. ⁵Department of Physics, Caltech, Pasadena, California 91125, USA; Institute for Astrophysics, University of Minnesota, Minneapolis, Minnesota 55405, USA. ⁶Department of Physics, New Mexico Institute of Mining and Technology, Socorro, New Mexico 85723, USA. ⁷National Radio Astronomy Observatory, 1003 Lopezville Road, Socorro, New Mexico 85723, USA. ⁸Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.

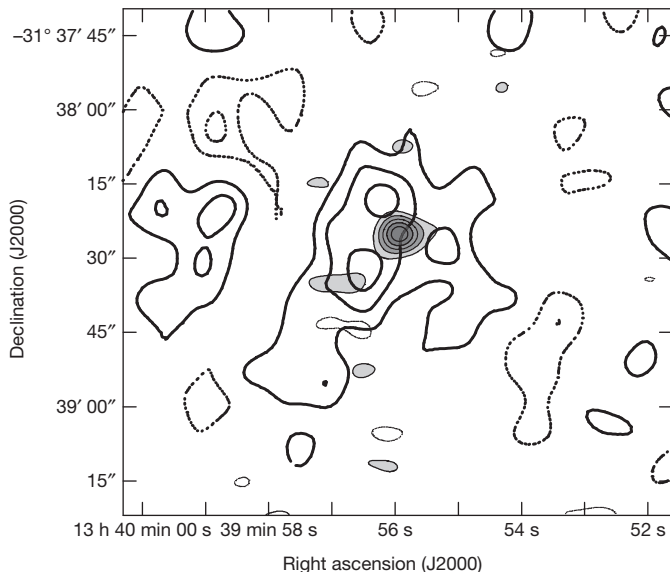


Figure 2 | Dust and gas in NGC 5253. SMA image of continuum dust emission at 870 μm (greyscale), with 350- μm dust continuum emission from SHARC at the Caltech Submillimeter Observatory (contours) superimposed. The SMA continuum image has been smoothed to 6'' resolution to show emission from the streamer. The SHARC image has been smoothed to 12.7''; contours are 2σ . Coordinates of the SHARC image are uncertain to $\sim 5''$ (see Methods).

We assume that the linewidth is gravitational, with no winds or outflow; this is therefore an upper limit to M_{vir} . The virial mass includes both gas and stars, but we can constrain the stellar mass. The mass in stars exciting the supernova, M_{stars} , can be predicted from the Lyman continuum rate of $N_{\text{Ly}\alpha} = (7 \pm 2) \times 10^{52} \text{ s}^{-1}$ (ref. 6) and Bracket γ equivalent width of 255 Å (ref. 10). The star cluster has mass $M_{\text{stars}} = 1.1^{+0.7}_{-0.2} \times 10^6 M_{\odot}$ (see Methods). We then obtain a gas mass $M_{\text{gas}} = M_{\text{vir}} - M_{\text{stars}} = (7 \pm 4) \times 10^5 M_{\odot}$ if the cluster is embedded in the cloud. This treatment assumes that the CO kinematics trace all of the cloud and that there is no extensive layer of H_2 without CO; however, the dust continuum and CO sizes are nearly identical, consistent with the dust and gas mass being contained within Cloud D.

Other methods to estimate H_2 mass are problematic for Cloud D. $\text{CO}(3 \rightarrow 2)$ is optically thin, but the excitation temperature is not determined, nor is CO/H_2 known. The line strength of $41 \pm 8 \text{ Jy km s}^{-1}$ gives $M_{\text{CO}} = (4 \pm 1) M_{\odot} (T/200 \text{ K})$. For a Galactic abundance ratio of $[\text{CO}]/[\text{H}_2] = 8.5 \times 10^{-5}$, the H_2 mass would be $M_{\text{H}_2} = 5 \times 10^4 M_{\odot} (T/200 \text{ K})$. This is one-tenth of the value for M_{gas} derived above. Intense radiation fields and high temperature will affect the chemistry and the relative abundance of CO.

Dust continuum emission can trace gas mass, but it is also unreliable in Cloud D. The 870- μm continuum flux density for Cloud D is $S_{870\mu\text{m}} = 72 \pm 10 \text{ mJy}$, consistent with previous measurements¹¹, of which free-free emission⁸ contributes $S_{870\mu\text{m}}^{\text{ff}} = 38 \pm 4 \text{ mJy}$, leaving dust emission $S_{870\mu\text{m}}^{\text{dust}} = 34 \pm 14 \text{ mJy}$. Our 350- μm image gives $S_{350\mu\text{m}}^{\text{dust}} = 1.0 \pm 0.2 \text{ Jy}$ for Cloud D, consistent with the 870- μm flux. Adopting the dust opacity of the Large Magellanic Cloud¹² (see Methods), we find an observed dust mass, $M_{\text{dust}} = (1.5 \pm 0.2) \times 10^6 M_{\odot} (T/45 \text{ K})^{-1}$. To obtain gas mass, we need a gas-to-dust ratio (GTD). Scaled as the oxygen abundance of NGC 5253, 0.2–0.3 solar^{13,14}, $\text{GTD} \approx 650$, within the range 340–1,200 inferred for the Magellanic clouds^{15–17}. The observed dust mass is five-fold the $\sim 3,000 M_{\odot}$ of dust expected for a $2 \times 10^6 M_{\odot}$ gas cloud at $\text{GTD} = 650$. If instead we compare the observed dust mass with the dynamical estimate of gas mass, we derive $\text{GTD} \approx 47$ for the embedded cluster, and in the unlikely case that the cluster is not embedded, $\text{GTD} \approx 120$. Dust is an expected result of mass loss from massive, short-lived stars. Stellar models indicate that a cluster of massive stars of age

4.4 Myr, consistent with recombination line equivalent widths, will expel 20,000–30,000 M_{\odot} of elements carbon, oxygen, silicon, magnesium and iron depending on the cluster mass and initial mass function, of which ~ 30 –50% will be in the form of dust (see Methods). To produce the amounts of dust and ionizing photons observed, given the upper limit imposed by the dynamical mass, suggests that the stellar initial mass function is top-heavy, with a lower mass cutoff of at least 2 – $3 M_{\odot}$ (see Methods.) The star cluster has probably produced most of the dust. To infer a gas mass on the basis of the observed dust emission for Cloud D from a GTD scaled to the global metallicity of NGC 5253 without accounting for *in situ* dust production, as has been done for other galaxies¹⁸, would give an erroneously high gas mass and underestimate the star-formation efficiency.

Given the peculiarities of Cloud D, the most reliable gas mass is dynamical. We use this mass to calculate the star-formation efficiency, η . A lower limit to η occurs if the gravitational mass is all gas—no stars—so that $\eta = 1.1^{+0.7}_{-0.2} \times 10^6 M_{\odot} / (1.1^{+0.7}_{-0.2} + 1.8^{+0.2}_{-0.7}) \times 10^6 M_{\odot} \approx 38^{+24}_{-7}\%$. Even in this case, η significantly exceeds the $\eta < 1\%$ of Galactic giant molecular clouds and the highest efficiencies of $\eta \approx 15$ –20% seen in individual cloud cores¹⁹ in the Galaxy. However, it is almost certain that the star cluster is located within the cloud, given the subarcsecond positional coincidence of nebular emission and CO, the precise kinematic coincidence of nebular $\text{H}53\alpha$ (ref. 8) and CO line centroids, and the high extinction to the cluster. Thus the stellar mass also contributes to the linewidth. For the more realistic case that the cluster is embedded within Cloud D, $\eta = (1.1^{+0.7}_{-0.2}) \times 10^6 M_{\odot} / 1.8^{+0.2}_{-0.7} \times 10^6 M_{\odot} = 61^{+84}_{-16}\%$. This value exceeds even the canonical $\sim 50\%$ (see Methods) needed to allow a star cluster to survive in its current bound state with rapid gas dispersal. If dust competes with gas for ultraviolet photons, the Lyman continuum rate and stellar mass have been underestimated, and η is even higher. If there are winds or outflows contributing to the CO linewidth, η is higher. The large dust mass favours larger-mass cluster models for which η is higher. These values for η are uncertain, but they are free of the systematics due to standard assumptions such as gas-to-dust ratio, relative CO abundance or CO conversion factor. The star-formation efficiency of Cloud D is unusually high, implying gas consumption timescales of ~ 10 Myr.

A measurement of star-formation efficiency is a snapshot in time; η could be high because the gas has been incorporated into stars or because the young stars have already dispersed the gas. How and when a young star cluster disperses its gas is crucial to its survival^{2,20}. For Cloud D, gas dispersal models are strongly constrained by the youth of the embedded star cluster¹⁰, its positional coincidence with the cloud, lack of evidence for supernovae²¹, and small $\text{CO}(3 \rightarrow 2)$ linewidth. Apparently not much has yet escaped this cloud.

Cloud D is a strange molecular cloud: hot, dusty, and small in mass relative to its young star cluster. It is found in a dark-matter-dominated galaxy. Its unusual properties may indicate a mode of star formation different from that observed in disk galaxies, including luminous infrared galaxies. Models of stochastic star formation for turbulently supported giant molecular clouds in our Galaxy suggest that star-formation efficiencies are 1% in a free-fall time²², which implies that the ultimate efficiency can be limited if star formation is quenched by massive stellar feedback. An extended period of star formation might be facilitated if Cloud D is compressed by an external influence, as for example by a streamer of gas force-fed into the star-forming region by the galactic potential. Our data in NGC 5253 could support such a model. The streamer contains $\sim 2 \times 10^6 M_{\odot}$ of gas extending ~ 200 –300 pc along the minor axis, entering the galaxy at a rate of $\sim 20 \text{ pc Myr}^{-1}$. The streamer can fuel star formation at the present rate of ~ 0.1 – $0.2 M_{\odot} \text{ yr}^{-1}$ for the next 10 Myr. This dwarf spheroidal galaxy is not rotationally supported²³; multiple accreting streams from its extensive H I halo^{24,25} could be responsible for its global dynamics and morphology²⁶ as well as its spheroidal system of massive star clusters spanning billions of years in age^{27–29}. NGC 5253 may illustrate a new mode of highly efficient star cluster formation triggered by cold-stream accretion³⁰.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 28 July 2014; accepted 6 January 2015.

1. Kennicutt, R. C. Jr. The global Schmidt Law in star-forming galaxies. *Astrophys. J.* **498**, 541–552 (1998).
2. Bastian, N. & Goodwin, S. P. Evidence for the strong effect of gas removal on the internal structure of young stellar clusters. *Mon. Not. R. Astron. Soc.* **369**, L9–L13 (2006).
3. Meier, D. S., Turner, J. L. & Beck, S. C. Molecular gas and the young starburst in NGC 5253 revisited. *Astron. J.* **124**, 877–885 (2002).
4. Turner, J. L., Beck, S. C. & Ho, P. T. P. The radio supernova in NGC 5253. *Astrophys. J.* **532**, L109–L112 (2000).
5. Turner, J. L., Beck, S. C. & Hurt, R. L. A CO map of the dwarf starburst galaxy NGC 5253. *Astrophys. J.* **474**, L11–L14 (1997).
6. van der Tak, F. F. S., Black, J. H., Schöier, F. L., Jansen, D. J. & van Dishoeck, E. F. A computer program for fast non-LTE analysis of interstellar line spectra. With diagnostic plots to interpret the line intensity ratios. *Astron. Astrophys.* **468**, 627–635 (2007).
7. Cresci, G., Vanzi, L., Sauvage, M., Santagelo, G. & van der Werf, P. Integral-field near-infrared spectroscopy of two blue dwarf galaxies: NGC 5253 and He 2–10. *Astron. Astrophys.* **520**, A82 (2010).
8. Rodríguez-Rico, C. A., Goss, W. M., Turner, J. L. & Gómez, Y. VLA H53 α observations of the central region of the super star cluster galaxy NGC 5253. *Astrophys. J.* **670**, 295–300 (2007).
9. Gray, W. J. & Scannapieco, E. Thermal and chemical evolution of collapsing filaments. *Astrophys. J.* **768**, 174 (2013).
10. Alonso-Herrero, A. *et al.* Obscured star formation in the central region of the dwarf galaxy NGC 5253. *Astrophys. J.* **612**, 222–237 (2004).
11. Hirashita, H. Properties of free-free, dust and CO emissions in the starbursts of blue compact dwarf galaxies. *Mon. Not. R. Astron. Soc.* **429**, 3390–3401 (2013).
12. Galliano, F. *et al.* Non-standard grain properties, dark gas reservoir, and extended submillimeter excess, probed by Herschel in the Large Magellanic Cloud. *Astron. Astrophys.* **536**, A88 (2011).
13. Kobulnicky, H. A., Skillman, E. D., Roy, J.-R., Walsh, J. R. & Rosa, M. R. HST FOS spectroscopy of localized chemical enrichment from massive stars in NGC 5253. *Astrophys. J.* **477**, 679–692 (1997).
14. López-Sánchez, Á. R., Esteban, C., García-Rojas, J., Peimbert, M. & Rodríguez, M. The localized chemical pollution in NGC 5235 revisited: results from deep echelle spectrophotometry. *Astrophys. J.* **656**, 168–185 (2007).
15. Bot, C., Boulanger, F., Lagache, G., Cambrésy, L. & Egret, D. Multi-wavelength analysis of the dust emission in the Small Magellanic Cloud. *Astron. Astrophys.* **423**, 567–577 (2004).
16. Gordon, K. *et al.* The dust-to-gas ratio in the Small Magellanic Cloud tail. *Astrophys. J.* **690**, L76–L80 (2009).
17. Roman-Duval, J. *et al.* Dust/gas correlations from Herschel observations. *Astron. Astrophys.* **518**, L74 (2010).
18. Shi, Y. *et al.* Inefficient star formation in extremely metal poor galaxies. *Nature* **514**, 335–338 (2014).
19. Olmi, L. & Testi, L. Constraints on star formation theories from the Serpens molecular cloud and protocluster. *Astron. Astrophys.* **392**, 1053–1068 (2002).
20. Smith, R., Goodwin, S., Fellhauer, M. & Assmann, P. Infant mortality in the hierarchical merging scenario: dependence on gas expulsion time-scales. *Mon. Not. R. Astron. Soc.* **428**, 1303–1311 (2013).
21. Beck, S. C., Turner, J. L., Ho, P. T. P., Lacy, J. H. & Kelly, D. M. The central star cluster of the star-forming dwarf galaxy NGC 5253. *Astrophys. J.* **457**, 610–615 (1996).
22. Krumholz, M. R. & McKee, C. F. A general theory of turbulence-regulated star formation, from spirals to ultraluminous infrared galaxies. *Astrophys. J.* **630**, 250–268 (2005).
23. Caldwell, N. & Phillips, M. M. Star formation in NGC 5253. *Astrophys. J.* **338**, 789–803 (1989).
24. Kobulnicky, H. A. & Skillman, E. D. Inflows and outflows in the dwarf starburst galaxy NGC 5253: high-resolution HI observations. *Astron. J.* **135**, 527–537 (2008).
25. López-Sánchez, Á. R. *et al.* The intriguing HI gas in NGC 5253: an infall of a diffuse, low-metallicity HI cloud? *Mon. Not. R. Astron. Soc.* **419**, 1051–1069 (2012).
26. Cen, R. Evolution of cold streams and the emergence of the Hubble sequence. *Astrophys. J.* **789**, L21 (2014).
27. Harris, J., Calzetti, D., Gallagher, J. S., Smith, D. A. & Conselice, C. J. The recent cluster formation histories of NGC 5253 and NGC 3077: environmental impact on star formation. *Astrophys. J.* **603**, 503–522 (2004).
28. Cresci, G., Vanzi, L. & Sauvage, M. The star cluster population of NGC 5253. *Astron. Astrophys.* **433**, 447–454 (2005).
29. de Grijs, R., Anders, P., Zackrisson, E. & Göran, Ö. The NGC 5253 star cluster system. I. Standard modelling and infrared-excess. *Mon. Not. R. Astron. Soc.* **431**, 2917–2932 (2013).
30. Dekel, A. *et al.* Cold streams in early massive hot haloes as the main mode of galaxy formation. *Nature* **457**, 451–454 (2009).

Acknowledgements We thank J. Carpenter, S. Goodwin, M. Heyer, L. Hunt, R. Hurt, M. Jura, C. Lada, C. Leitherer and S. Van Dyk for assistance with the analysis. The Submillimeter Array is a joint project between the Smithsonian Astrophysical Observatory and the Academia Sinica Institute of Astronomy and Astrophysics and is funded by the Smithsonian Institution and the Academia Sinica.

Author Contributions J.L.T., S.C.B., D.J.B., A.K. and D.S.M. performed the observations. J.L.T., S.C.B. and P.T.P.H. conceived the project and wrote the observing proposal. J.-H.Z. reduced and imaged the Submillimeter Array 870- μ m data; A.K. reduced, imaged and analysed the SHARC 350- μ m data. J.L.T. and S.M.C. obtained derived quantities and performed data analysis. J.L.T. wrote the first draft and constructed figures. All authors read, discussed and commented on the draft.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.L.T. (turner@astro.ucla.edu).

METHODS

Submillimeter Array observations. NGC 5253 was observed with the Submillimeter Array (SMA)³¹ on 2011 April 17. The observing frequency was $\nu_{\text{LO}} = 340.323$ GHz with 48 adjacent spectral windows covering 4 GHz bandwidth for each of two sidebands. The CO $J = 3 \rightarrow 2$ rotational transition at $\nu_0 = 345.79599$ GHz was in the upper sideband. The array was in the subcompact configuration covering the visibility baselines between 9 and 80 k λ , corresponding to the angular scales between 29'' and 2''. The phase centre was $\alpha_{J2000} = 13 \text{ h } 39 \text{ min } 56.249 \text{ s}$, $\delta_{J2000} = -31^\circ 38' 29.00''$. Calibration and reduction were performed with MIRIAD³². The instrumental bandpass was corrected using the quasar 3C 279; complex gains were calibrated using the nearby quasar J1316–336; the flux density scale was determined from the planet model of Neptune. Continuum and line emission were separated using the task UVLIN by fitting a linear model to line-free channels. The result is a cube of 25 channels of 10 km s^{-1} and a continuum map with an effective bandwidth of 8 GHz, convolved to a beam $4'' \times 2''$, position angle = 0° , shown in Extended Data Fig. 1. Final noise levels were 3 mJy per beam in the continuum map, and 50 mJy per beam in the individual 10 km s^{-1} channels.

CSO SHARC observations. The 350- μm continuum observations were made with the SHARC camera³³ at the Caltech Submillimeter Observatory on 1999 January 11–12, with 225-GHz opacities around 0.035 and 0.075 for the respective dates. The data consist of 2.2 h of on-the-fly mapping with a $\sim 60''$ chopping secondary at 4.132 Hz, and were reduced with CRUSH³⁴, using an enhanced implementation of the Emerson II deconvolution algorithm³⁵, which uses sky rotation to fill in the poorly sampled spatial frequencies of the dual-beam chop. CRUSH removes direct-current detector offset and correlated sky-noise residuals; flatfields detectors based on sky response; and performs noise weighting, whitening and despiking. The main beam is $9''$ full width at half-maximum at 350 μm , but the image presented here was smoothed to 12.7'' resolution. From observations of Mars taken immediately before the SHARC observations, at a similar elevation, it is estimated that the pointing is good to $\sim 5''$ root mean squared (r.m.s.). The systematic aperture flux calibration of the 350- μm image is estimated to be good to within 7% r.m.s.

Relation of NGC 5253 to M83 and distance. NGC 5253 is a dwarf spheroidal galaxy of the Cen A/M83 galaxy complex³⁶, with a stellar mass of $\sim 1.5 \times 10^8 M_\odot$ (ref. 37) and an estimated³⁸ total mass, including dark matter, about tenfold higher. It is close to the large spiral galaxy M83 in projection. However, the distance to M83, 4.8 Mpc (ref. 39), is significantly larger than the distance to NGC 5253, at 3.8 Mpc (ref. 40). The H I streamer system^{24,25} in the halo of NGC 5253, from which the CO streamer seems to emanate, strongly suggests that this dwarf galaxy has had some encounter in its past, but M83 does not seem to be responsible.

Cloud D CO emission. The $J = 3$ level of CO corresponds to an energy E_u/k of 33 K, a temperature that begins to distinguish actively star-forming clumps from giant molecular clouds. Cloud D is bright in CO($3 \rightarrow 2$), but only weakly detected in CO($2 \rightarrow 1$) (ref. 3), and not at all in CO($1 \rightarrow 0$) (ref. 5). The total flux of CO($3 \rightarrow 2$) emission in the galaxy and streamer is $110 \pm 20 \text{ Jy km s}^{-1}$, about 30% less than the single-dish flux⁴¹. This is a typical value for local galaxies, because the array configuration is insensitive to structures $< 30''$ in extent; the value is consistent with the extended streamer emission and with the JCMT-SCUBA continuum image⁴². The CO($2 \rightarrow 1$) image is shown in Extended Data Fig. 2, overlaid on the SMA CO($3 \rightarrow 2$) image. CO($3 \rightarrow 2$) was not detected in previous SMA observations¹¹ because of insufficient signal-to-noise. Located at $\alpha_{J2000} = 13 \text{ h } 39 \text{ min } 55.943 \text{ s} \pm 0.003 \text{ s}$, $\delta_{J2000} = -31^\circ 38' 25.097'' \pm 0.05''$, the Cloud D CO($3 \rightarrow 2$) source is coincident to within $\pm 0.5''$ with the core of the supernova as defined by high-brightness 7-mm free-free emission⁴³. The CO($3 \rightarrow 2$) line centre is at a heliocentric velocity of $397.5 \pm 0.6 \text{ km s}^{-1}$ and the CO flux of Cloud D is 41 Jy km s^{-1} . The size of the CO source in the integrated intensity map deconvolved from the beam is $2.8'' \times 1.5'' \pm 7\%$, position angle $12^\circ \pm 1^\circ$. The slight northward extension is consistent with features seen^{4,8,21,43,44} in free-free emission, but it is also in the same direction as the elongation of the beams for northern synthesis arrays for this source.

Cloud D virial mass. The width of the CO($3 \rightarrow 2$) line is $\sigma = 9.2 \pm 0.6 \text{ km s}^{-1}$, based on a least-squares fit to a Gaussian line profile. We adopt a value for the radius of half the full width at half-maximum of the geometric mean of the deconvolved source size, using $M_{\text{vir}} = \alpha M_\odot v^2 r$, where $v = 2.35\sigma$ (in km s^{-1}) and r is in parsecs, with coefficients of $\alpha = 190$, for $\rho \approx r^{-1}$, adopted here, and $\alpha = 126$ for $\rho \approx r^{-2}$ and 210 for $\rho \approx r^0$ giving the uncertainty limits in M_{vir} (ref. 45). We assume that the cloud is turbulently supported against gravity and dispersion-dominated, as for Galactic giant molecular clouds^{46,47}; inclination effects should therefore not be important. If Cloud D is not bound, or has flows that are super-gravitational, our estimate for the virial mass is an overestimate.

Supernova stellar mass. The stellar mass is based on STARBURST99 (refs 48, 49) modelling with the following constraints. First, the 7-mm flux density of the supernova is $47 \pm 4 \text{ mJy}$ for the central $2''$ (refs 8, 43). From this we obtain a Lyman continuum rate of $N_{\text{Ly}\alpha} = 7 \times 10^{52} \text{ s}^{-1}$ for a nebula at 12,000 K (refs 50, 51). Second, the cluster age must be consistent with Brackett γ equivalent width of 255 Å

(refs 10, 52) and mid-infrared ionic line ratios⁵³. Third, the cluster must be old enough to have Wolf–Rayet (WR) stars, which would explain the WR spectral signatures^{13,14,50,51,54,55}. Fourth, the cluster must have mass less than the virial mass of $1.8 \times 10^6 M_\odot$, which is also the maximum mass allowed by the [SIV] and Brackett α linewidths^{56,57}. We considered both Geneva high and Geneva $v = 40\%$ breakup stellar models; the high-velocity models permit older cluster ages. The main parameter to vary in fitting the models is the lower mass cutoff to initial mass function (IMF). We consider cluster models with standard Kroupa IMFs with upper mass cutoffs of $150 M_\odot$, and top-heavy IMFs. Kroupa IMFs with stellar masses down to $0.1 M_\odot$ cannot give cluster ages sufficiently old to have both WR stars and the given $N_{\text{Ly}\alpha}$ in view of the upper limit on the cluster mass. The IMF must be top-heavy: a cluster of $M_{\text{stars}} = 1.1^{+0.7}_{-0.2} \times 10^6 M_\odot$ requires a lower mass cutoff of $> 3 M_\odot$. This cluster mass is consistent with previous estimates^{58,59}. The Lyman continuum rate inferred from free-free emission may be less than the true value because of leakage from the H II region. In addition, dust can absorb as much as 50% of the ionizing photons in dense Galactic H II regions⁶⁰; this could also increase the stellar mass. Studies of the extended ionized gas in NGC 5253 (ref. 61) indicate a total galactic star formation rate of twice what we calculate for the supernova, but this could be ultraviolet photons from nearby slightly older clusters^{27–29,58}.

Cloud D continuum and dust mass. The strong 870- μm continuum source towards Cloud D consists of equal parts free-free emission from the H II region and dust emission. The continuum source is located at $\alpha_{J2000} = 13 \text{ h } 39 \text{ min } 55.948 \text{ s} \pm 0.005 \text{ s}$, $\delta_{J2000} = -31^\circ 38' 24.88'' \pm 0.11''$ (J2000). The 870- μm peak is $0.5''$ north and $0.32''$ west of the 7-mm continuum supernova core⁴³, which is within the uncertainties of our $4'' \times 2''$ beam. The continuum source agrees in position and size with the CO source. The total 870- μm flux density of Cloud D is $72 \pm 10 \text{ mJy}$. This flux is consistent with previous observations¹¹, and constitutes $\sim 40\%$ of the total 870- μm flux of 192 mJy for the galaxy as determined from a JCMT/SCUBA map⁴². We extrapolate a free-free flux density from the 7-mm value of $47 \pm 4 \text{ mJy}$ (refs 8, 43), using the $S_\nu \propto \nu^{-0.1}$ spectrum of optically thin emission, giving $S_{870\mu\text{m}}^{\text{ff}} = 38 \pm 4 \text{ mJy}$ for Cloud D. The dust emission is then $S_{870\mu\text{m}}^{\text{dust}} = S_{870\mu\text{m}} - S_{870\mu\text{m}}^{\text{ff}} = 34 \pm 14 \text{ mJy}$. This value is just consistent with the upper limit set at 1.3 mm (ref. 3). The SHARC flux is $S_{350\mu\text{m}}^{\text{dust}} = 3.7 \pm 0.5 \text{ Jy}$, of which $\sim 1.0 \pm 0.2 \text{ Jy}$ originates in Cloud D.

The dust mass is dependent on the submillimetre dust opacity and dust temperature. We adopt the opacity of the Large Magellanic Cloud¹², extrapolating from $\kappa(160 \mu\text{m}) = 16 \text{ cm}^2 \text{ g}^{-1}$ and $\beta = 1.7$ to obtain $\kappa(870 \mu\text{m}) = 0.9 \text{ cm}^2 \text{ g}^{-1}$. The 350- μm and 870- μm fluxes for Cloud D are consistent with $\beta = 1.7$. For dust temperature we adopt $T_{\text{dust}} = 45 \text{ K}$ based on IRAS Point Source Catalog fluxes⁶²; $T_{\text{dust}} < 57 \text{ K}$ based on IRAS 60- μm flux. These values of temperature and opacity give $M_{\text{dust}} = (1.5 \pm 0.1) \times 10^4 M_\odot (T/45 \text{ K})^{-1}$ for the dust mass of Cloud D, with the uncertainty based on the flux. Previous determinations of the dust mass in NGC 5253 were for the entire galaxy, including the streamer, based on the large-aperture JCMT/SCUBA⁴² flux density; if scaled to our flux of 34 mJy for Cloud D only, these models^{59,63} would give a dust mass of $M_{\text{dust}} \approx (2\text{--}3) \times 10^4 M_\odot$, consistent with a cooler dust temperature.

Dust yield from the cluster and GTD. The GTD estimated from our dust mass of $1.5 \times 10^4 M_\odot$ and our gas mass derived from virial and stellar masses is ~ 47 (embedded cluster) or 120 (cluster outside cloud). Either is significantly lower than the value of 650 predicted⁶⁴ from scaling the Galactic value of 160 (ref. 65) to the metallicity of NGC 5253. We argue that the high dust mass is from *in situ* enrichment by the cluster. The Brackett γ equivalent width of 255 Å (refs 10, 52) indicates a cluster age consistent with the presence of WR activity, so mass loss is expected. If the original progenitor cloud had a mass of $\sim 2 \times 10^6 M_\odot$, there would initially have been $\sim 3,000 M_\odot$ of dust in Cloud D, one-fifth of the amount that we observe. From STARBURST99 models with $z = 0.004$, Geneva high-mass-loss stellar models, for a $1.1 \times 10^6 M_\odot$ cluster of age 4.4 Myr with an initial mass function of range 3– $150 M_\odot$ and $\log N_{\text{Ly}\alpha} = 52.84$, one would expect a yield in elements C, N, O, Mg, Si and Fe of $\sim 24,000 M_\odot$, of which an estimated 30–50% would be in the form of dust^{65,66}. If the cluster is more massive or more top-heavy, the mass loss can be higher. The Geneva rotating models give yields that are lower but still sufficient to explain the dust mass. The optical spectrum of NGC 5253 does not yet reflect this localized enrichment: the metallicity of ~ 0.25 solar is based on nebular lines^{13,14} from the extended nuclear H II region, whereas the embedded star cluster is behind at least 16 magnitudes of extinction. However, the small CO linewidth suggests that what has been produced in the cloud has so far largely stayed in the cloud.

Cloud D excitation modelling. The CO($3 \rightarrow 2$) to CO($2 \rightarrow 1$) line ratio can be used to constrain gas density and temperature. Extended Data Fig. 2 demonstrates clear differences between the emission in Cloud D and the streamer. RADEX⁶ was used to perform non-LTE modelling of the line ratios. We assumed a black-body radiation field and a column density of 10^{16} cm^{-2} , consistent with the observed brightness. Allowed values of density and kinetic temperature for the adopted value of $I_{32}/I_{21} = 2.25$ and the 1σ lower limit of 2.0 are shown in Extended Data Fig. 3. For the value of 2.25, the lower limit on the gas kinetic temperature is $T_K \approx 350 \text{ K}$; however,

the ratio of 2.1 is within the uncertainties and would give $T_K > 200$ K. The indication of thermal ratios in the near-infrared H_2 line ratios⁷ would also support a high temperature for this cloud.

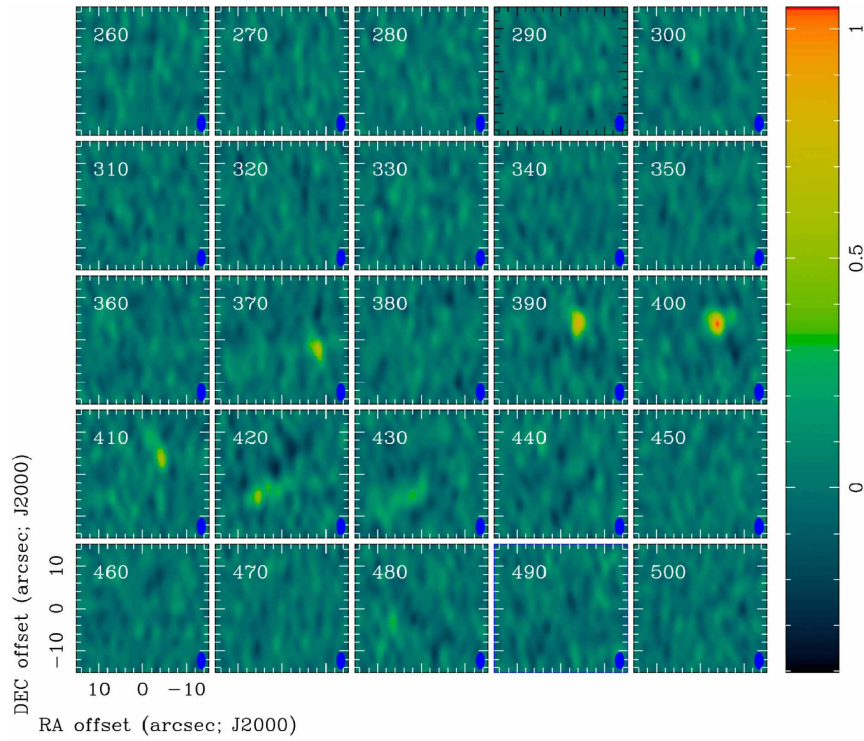
Mass of Cloud D based on X_{CO} . X_{CO} masses are based on $CO(1\rightarrow 0)$ emission, which has not been detected in Cloud D. Given that the CO emission is optically thin, it seems that the X_{CO} value³ would underpredict the H_2 mass by a factor of ~ 8 . **Streamer kinematics.** The streamer has been detected previously in $CO(1\rightarrow 0)$ and $CO(2\rightarrow 1)$ (refs 3, 5). The emission is found at heliocentric velocities of 410–430 km s^{-1} , which is red-shifted by about 20 km s^{-1} with respect to the galaxy and the supernova Cloud D. High-resolution VLA images²⁴ also detect H I emission coincident with this streamer. The streamer coincides with filamentary emission in nebular lines of oxygen⁶⁷ and sulphur⁶⁸; it has been suggested that this is an ‘ionization cone’, possibly even due to an active galactic nucleus⁶⁸. We suggest that this emission is due to leakage of photons from the starburst, which are ionizing the surface of the infalling streamer.

Streamer CO and dust properties and GTD. The $CO(3\rightarrow 2)$ emission originates largely from a single cloud, Cloud ‘C’ (ref. 3). For its line strength of $I_{32} = 3.9 \pm 0.6 \text{ K km s}^{-1}$, the ratio I_{32}/I_{21} is 1.0 ± 0.3 . RADEX models of the ratio are consistent with optically thick and cooler gas, $T \approx 20$ K, with $n \approx 10^{3.5}–10^4 \text{ cm}^{-3}$ (Extended Data Fig. 3). The molecular mass of the streamer based on $CO(2\rightarrow 1)$ (ref. 3) is $M_{H_2} = 2 \times 10^6 M_\odot$ for a Galactic conversion factor, $X_{CO} = 2 \times 10^{20} \text{ cm}^{-2} (\text{K km s}^{-1})^{-1}$. The virial mass is $3 \times 10^6 M_\odot$ (ref. 3). The 870- μm dust continuum emission follows the CO (Fig. 2). The 870- μm flux density of the streamer is $26 \pm 8 \text{ mJy}$, which is all dust (Fig. 2). Adopting the dust opacity $\kappa(870 \mu\text{m}) = 0.9 \text{ cm}^2 \text{ g}^{-1}$ and dust temperature $T_d = 20$ K, we obtain a dust mass of $M_{\text{dust}} = 2.6 \times 10^4 M_\odot$. The streamer thus has GTD ≈ 120 using the virial mass for the gas mass. If the H I gas²⁴ is added, the total H + H_2 mass becomes $4.3 \times 10^6 M_\odot$, which gives GTD = 170. That the streamer is molecular gas, and—even more surprisingly—dense molecular gas, is difficult to understand. Molecular gas favours high-pressure environments⁶⁹ such as the midplanes of the central regions of spiral disks. Even though the filament seems to be in a low-pressure environment, it is not only molecular, but also dense. Models of the streamer as an example of a primordial cooling filament, in which the gas collapses towards the centre of the dark-matter potential, are able to produce the observed inflow rate of gas of $\sim 0.1–0.2 M_\odot \text{ yr}^{-1}$ but are unable to reproduce the formation of the observed giant molecular clouds⁷. The streamer may be previously enriched gas.

Cloud F. Cloud F, located about 5'' to the southwest of Cloud D, was not detected in previous CO observations^{3,5}. Clouds D and F are the only two detected giant molecular clouds within NGC 5253 proper. Using the Galactic CO conversion factor and assuming optically thick emission, for the observed flux of $17 \pm 6 \text{ Jy km s}^{-1}$ we obtain a mass of $M_{H_2} \approx 4 \times 10^5 M_\odot$ for Cloud F.

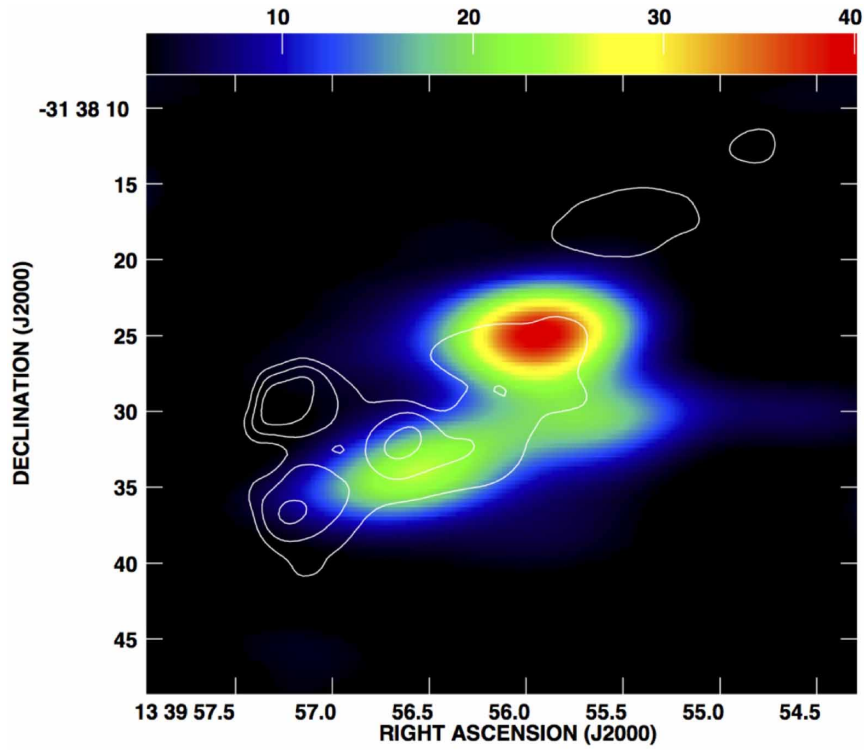
Star-formation efficiency and cluster survival. The canonical value of $\eta = 50\%$ is from virial considerations for the survival of a bound cluster with mass loss on timescales less than the crossing time⁷⁰. It is possible for a cluster to survive with lower efficiency, to $\sim 30\%$, if the gas is lost slowly and the cluster expands^{2,71}.

31. Moran, J. M. & Ho, P. T. P. Smithsonian Submillimeter Wavelength Array. *Proc. SPIE* **2200**, 335–346 (1994).
32. Sault, R. J., Teuben, P. J. & Wright, M. C. H. in *Astronomical Data Analysis Software and Systems IV* (eds Shaw, R., Payne, H. E. & Hayes, J. J. E.) 433–436 (Astron. Soc. Pacif. Conference Series Vol. 77, 1995).
33. Wang, N. *et al.* A submillimeter high angular resolution bolometer array camera for the Caltech submillimeter observatory. *Proc. STT* **7**, 426–438 (1996).
34. Kovács, A. CRUSH: fast and scalable data reduction for imaging arrays. *Proc. SPIE* **7420**, 45–59 (2008).
35. Emerson, D. T. in *Multi-feed Systems for Radio Telescopes* (eds Emerson, D. T. & Payne, J. M.) 309–317 (Astron. Soc. Pacif. Conference Series Vol. 75, 1995).
36. Karachentsev, I. D. *et al.* The Hubble flow around the Centaurus A/M83 galaxy complex. *Astron. J.* **133**, 504–517 (2007).
37. Martin, C. L. The impact of star formation on the interstellar medium in dwarf galaxies. II. The formation of galactic winds. *Astrophys. J.* **506**, 222–252 (1998).
38. Persic, M., Salucci, P. & Stel, F. The universal rotation curve of spiral galaxies. I. The dark matter connection. *Mon. Not. R. Astron. Soc.* **281**, 27–47 (1996).
39. Radburn-Smith, D. J. *et al.* The GHOSTS survey. I. Hubble Space Telescope Advanced Camera for Surveys data. *Astrophys. J.* **195**, 18 (2011).
40. Sakai, S., Ferrarese, L., Kennicutt, R. C. Jr & Saha, A. The effect of metallicity on Cepheid-based distances. *Astrophys. J.* **608**, 42–61 (2004).
41. Meier, D. S., Turner, J. L., Crosthwaite, L. P. & Beck, S. C. Warm molecular gas in dwarf starburst galaxies: CO(3–2) observations. *Astron. J.* **121**, 740–752 (2001).
42. James, A., Dunne, L., Eales, S. & Edmunds, M. G. SCUBA observations of galaxies with metallicity measurements: a new method for determining the relation between submillimetre luminosity and dust mass. *Mon. Not. R. Astron. Soc.* **335**, 753–761 (2002).
43. Turner, J. L. & Beck, S. C. The birth of a super star cluster. *Astrophys. J.* **602**, L85–L88 (2004).
44. Turner, J. L., Ho, P. T. P. & Beck, S. C. The radio properties of NGC 5253 and its unusual HII regions. *Astron. J.* **116**, 1212–1220 (1998).
45. MacLauren, I., Richardson, K. M. & Wolfendale, A. W. Corrections to virial estimates of molecular cloud masses. *Astrophys. J.* **333**, 821–825 (1988).
46. Solomon, P. M., Rivolo, A. R., Barrett, J. & Yahil, A. Mass, luminosity, and line width relations of Galactic molecular clouds. *Astrophys. J.* **319**, 730–741 (1987).
47. Heyer, M. & Brunt, C. M. The universality of turbulence in Galactic molecular clouds. *Astrophys. J.* **615**, L45–L48 (2004).
48. Leitherer, C. *et al.* Starburst99: synthesis models for galaxies with active star formation. *Astrophys. J. Suppl.* **123**, 3–40 (1999).
49. Leitherer, C. *et al.* The effects of stellar rotation. II. A comprehensive set of STARBURST99 models. *Astrophys. J.* **712** (Suppl.), 18 (2014).
50. Walsh, J. R. & Roy, J.-R. Optical spectroscopic and abundance mapping of the amorphous galaxy NGC 5253. *Mon. Not. R. Astron. Soc.* **239**, 297–324 (1989).
51. Monreal-Ibero, A., Walsh, J. R. & Vilchez, J. M. The ionized gas in the central region of NGC 5253. 2D mapping of the physical and chemical properties. *Astron. Astrophys.* **544**, A60 (2012).
52. Davies, R. I., Sugai, H. & Ward, M. J. Star-forming regions in blue compact dwarf galaxies. *Mon. Not. R. Astron. Soc.* **295**, 43–54 (1998).
53. Martín-Hernández, N. L., Schaerer, D. & Sauvage, M. High spatial resolution mid-infrared spectroscopy of NGC 5253: the stellar content of the embedded super-star cluster. *Astron. Astrophys.* **429**, 449–467 (2005).
54. Schaerer, D., Contini, T., Kunth, D. & Meynet, G. Detection of Wolf-Rayet stars of WN and WC subtypes in super-star clusters of NGC 5253. *Astrophys. J.* **481**, L75–L79 (1997).
55. Westmoquette, M. S., James, B., Monreal-Ibero, A. & Walsh, J. R. Piecing together the puzzle of NGC 5253: abundances, kinematics, and WR stars. *Astron. Astrophys.* **550**, A88 (2013).
56. Turner, J. L. *et al.* An extragalactic supernova confined by gravity. *Nature* **423**, 621–623 (2003).
57. Beck, S. C. *et al.* [SIV] in the NGC 5253 supernova: ionized gas kinematics at high resolution. *Astrophys. J.* **755**, 59 (2012).
58. Calzetti, D. *et al.* Dust and recent star formation in the core of NGC 5253. *Astron. J.* **114**, 1834–1849 (1997).
59. Vanzi, L. & Sauvage, M. Dust and super star clusters in NGC 5253. *Astron. Astrophys.* **415**, 509–520 (2004).
60. McKee, C. F. & Williams, J. P. The luminosity function of OB associations in the Galaxy. *Astrophys. J.* **476**, 144–165 (1997).
61. Martin, C. L. & Kennicutt, R. C. Jr. Soft X-ray emission from NGC 5253 and the ionized interstellar medium. *Astrophys. J.* **447**, 171–183 (1995).
62. Thronson, H. A. & Telesco, C. M. Star formation in active dwarf galaxies. *Astrophys. J.* **311**, 98–112 (1986).
63. Hunt, L., Bianchi, S. & Maiolino, R. The optical-to-radio spectral energy distributions of low-metallicity blue compact dwarf galaxies. *Astron. Astrophys.* **434**, 849–866 (2005).
64. Rémy-Ruyer, A. *et al.* Gas-to-dust ratios in local galaxies over a 2 dex metallicity range. *Astron. Astrophys.* **562**, A31 (2014).
65. Zubko, V., Dwek, E. & Arendt, R. G. Interstellar dust models consistent with extinction, emission, and abundance constraints. *Astrophys. J. Suppl.* **152**, 211–249 (2004).
66. Jenkins, E. B. A unified representation of gas-phase element depletions in the interstellar medium. *Astrophys. J.* **700**, 1299–1348 (2009).
67. Graham, J. Filamentary structure in NGC 5253. *Publ. Astron. Soc. Pacif.* **93**, 552–553 (1981).
68. Zastrow, J., Oey, M. S., Veilleux, S., McDonald, M. & Martin, C. L. An ionization cone in the dwarf starburst galaxy NGC 5253. *Astrophys. J.* **741**, L17 (2011).
69. Elmegreen, B. A pressure and metallicity dependence for molecular cloud correlations and the calibration of mass. *Astrophys. J.* **338**, 178–196 (1989).
70. Mathieu, R. D. Dynamical constraints on star formation efficiency. *Astrophys. J.* **267**, L97–L101 (1983).
71. Baumgardt, H., Kroupa, P. & Parmentier, G. The influence of residual gas expulsion on the evolution of the Galactic globular cluster system and the origin of the Population II halo. *Mon. Not. R. Astron. Soc.* **384**, 1231–1241 (2008).



Extended Data Figure 1 | Channel maps of CO(3→2) emission in NGC 5253. Positions are relative to a reference position 13 h 39 m 56.249 s, $-31^{\circ} 38' 29''$ (J2000). Channels are 10 km s^{-1} wide; the heliocentric velocity is

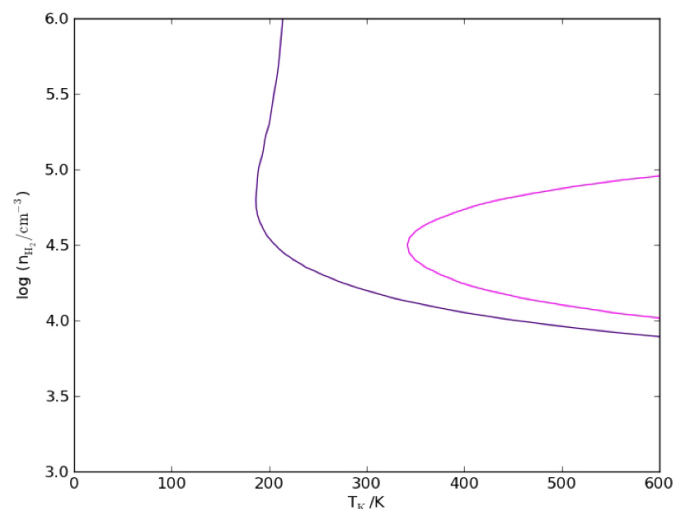
noted on the individual maps. The colour bar range maximum is 1 Jy per beam for each 10 km s^{-1} channel. The beam is $4'' \times 2''$, p.a. 0° .



Extended Data Figure 2 | SMA image of CO(3→2) with CO(2→1).

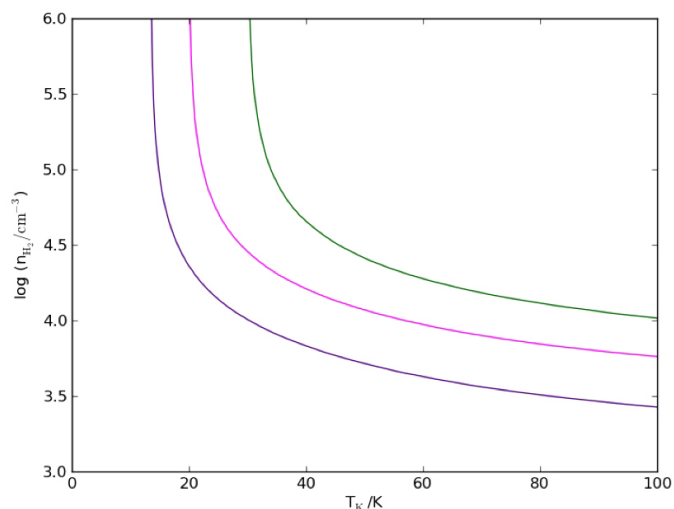
CO(3→2) emission is shown in colour, with an image² from the Owens Valley Millimeter Array of CO(2→1) in contours. The SMA CO(3→2) image has been smoothed from its original $4'' \times 2''$, p.a. 0° resolution to match the

$9.7'' \times 5''$, p.a. -84° , beam of the CO(2→1) image. The colour image flux range is $3\text{--}40 \text{ Jy km s}^{-1}$ per beam. Contours are linear multiples of 4 Jy km s^{-1} per beam.



Extended Data Figure 3 | RADEX modelling of Cloud D and streamer.

Escape probability transfer modelling of the CO(3 \rightarrow 2) to CO(2 \rightarrow 1) line ratio. Models were run using a black-body radiation field, spherical escape probability and $N_{\text{CO}} = 10^{16} \text{ cm}^{-2}$. **a**, Cloud D. Magenta is for the value



$I_{32}/I_{21} = 2.25$, the optically thin limit, and indigo is for the 1σ lower limit to the measured value of 2.6. There is no solution for $I_{32}/I_{21} = 2.6$. **b**, Streamer. Magenta is for $I_{32}/I_{21} = 0.98$, and indigo and green are solutions for the -1σ and $+1\sigma$ values, respectively.

Ligand-enabled *meta*-C–H activation using a transient mediator

Xiao-Chen Wang¹, Wei Gong¹, Li-Zhen Fang¹, Ru-Yi Zhu¹, Suhua Li¹, Keary M. Engle¹ & Jin-Quan Yu¹

Achieving site selectivity in C–H functionalization reactions is a significant challenge, especially when the target C–H bond is distant from existing functional groups^{1–5}. Coordination of a functional group to a metal is often a key driving force and control element in many important reactions including asymmetric hydrogenation⁶, epoxidation^{7,8} and lithiation⁹. Exploitation of this effect has led to the development of a broad range of directed C–H activation reactions^{10–14}. However, these C–H activation methods are limited to proximal C–H bonds, which are spatially and geometrically accessible from the directing functional group. The development of *meta*-selective C–H functionalizations remains a significant challenge^{1–5,15–17}. We recently developed a U-shaped template that can be used to overcome this constraint and have shown that it can be used to selectively activate remote *meta*-C–H bonds^{1,2}. Although this approach has proved to be applicable to various substrates and catalytic transformations^{3–5},

the need for a covalently attached, complex template is a substantial drawback for synthetic applications. Here we report an alternative approach employing norbornene as a transient mediator to achieve *meta*-selective C–H activation with a simple and common *ortho*-directing group. The use of a newly developed pyridine-based ligand is crucial for relaying the palladium catalyst to the *meta* position by norbornene after initial *ortho*-C–H activation. This catalytic reaction demonstrates the feasibility of switching *ortho*-selectivity to *meta*-selectivity in C–H activation of the same substrate by catalyst control.

Despite the recent reports of U-shaped templates to direct *meta*-C–H activation (Fig. 1a)^{1–5}, the development of more general and efficient approaches to achieve *meta*-C–H activation remains a significant task. Inspired by the unique reactivity of norbornene in palladium-catalysed reactions of aryl iodides (Fig. 1b)^{18–22}, we speculated that *ortho*-palladacycle I (Fig. 1c) could react with norbornene to provide an intermediate that

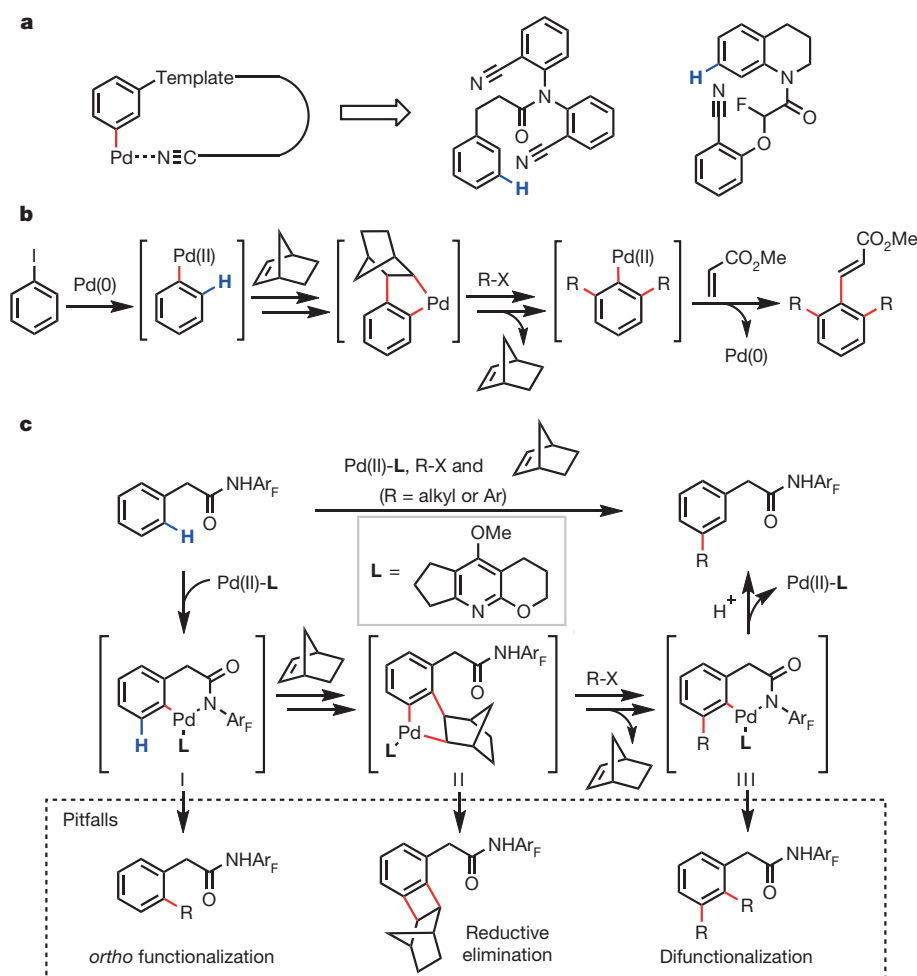


Figure 1 | Design of a new approach for *meta*-C–H activation. **a**, Covalent template strategy for remote C–H activation: previously reported *meta*-C–H activation directed by a U-shaped template. **b**, *ortho*-Alkylation and *ipso*-olefination in the Catellani reaction: the norbornene-mediated Catellani reaction functionalizes *ortho*-positions of aryl iodides. **c**, Norbornene as a transient mediator for *meta*-C–H activation: the development of a tailor-made ligand that switches the *ortho*-selectivity to *meta*-selectivity using norbornene as a transient mediator (this work). Ar_F, 4-(CF₃)C₆F₄. The activated C–H bonds are shown in blue and the newly formed bonds in red.

¹Department of Chemistry, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, California 92037, USA.

can undergo activation of the *meta*-C–H bond (intermediate **II**). Reaction of palladacycle **II** with a coupling partner would form a new C–C or C–heteroatom bond, and subsequent β -carbon elimination of norbornene followed by protodemetalation of the aryl–palladium bond would regenerate the palladium catalyst. This reaction pathway would therefore provide a new approach to achieving catalytic *meta*-selective C–H activation by using standard *ortho*-directing groups with norbornene as a transient mediator. We realized this concept through the design of a pyridine-based ligand that suppresses several undesired reaction pathways and promotes *meta*-selective alkylation and arylation of phenylacetic acid derivatives. We expect that this approach will enable many previously reported *ortho*-C–H activation reactions to be rendered highly *meta*-selective.

At the outset, we were aware that our proposed *meta*-C–H activation with norbornene as a transient mediator would face several key challenges (Fig. 1c). First, it was unclear whether 1,2-migratory insertion with norbornene could occur from a cyclopalladated intermediate (**I**) and whether this pathway would be faster than the competitive reaction of the palladacycle with a terminal electrophile. Nevertheless, we were encouraged by recent reports^{23,24} of C2-selective C–H alkylation of indoles via initial aminopalladation of norbornene with indolyl N–H bonds. Second, the resultant intermediate **II** could undergo reductive elimination to form a benzocyclobutene, thereby preventing the desired *meta*-C–H functionalization (Fig. 1c)²⁵. Third, rather than undergoing protodemetalation, intermediate **III** could be further functionalized to afford the undesired *ortho,meta*-difunctionalized products.

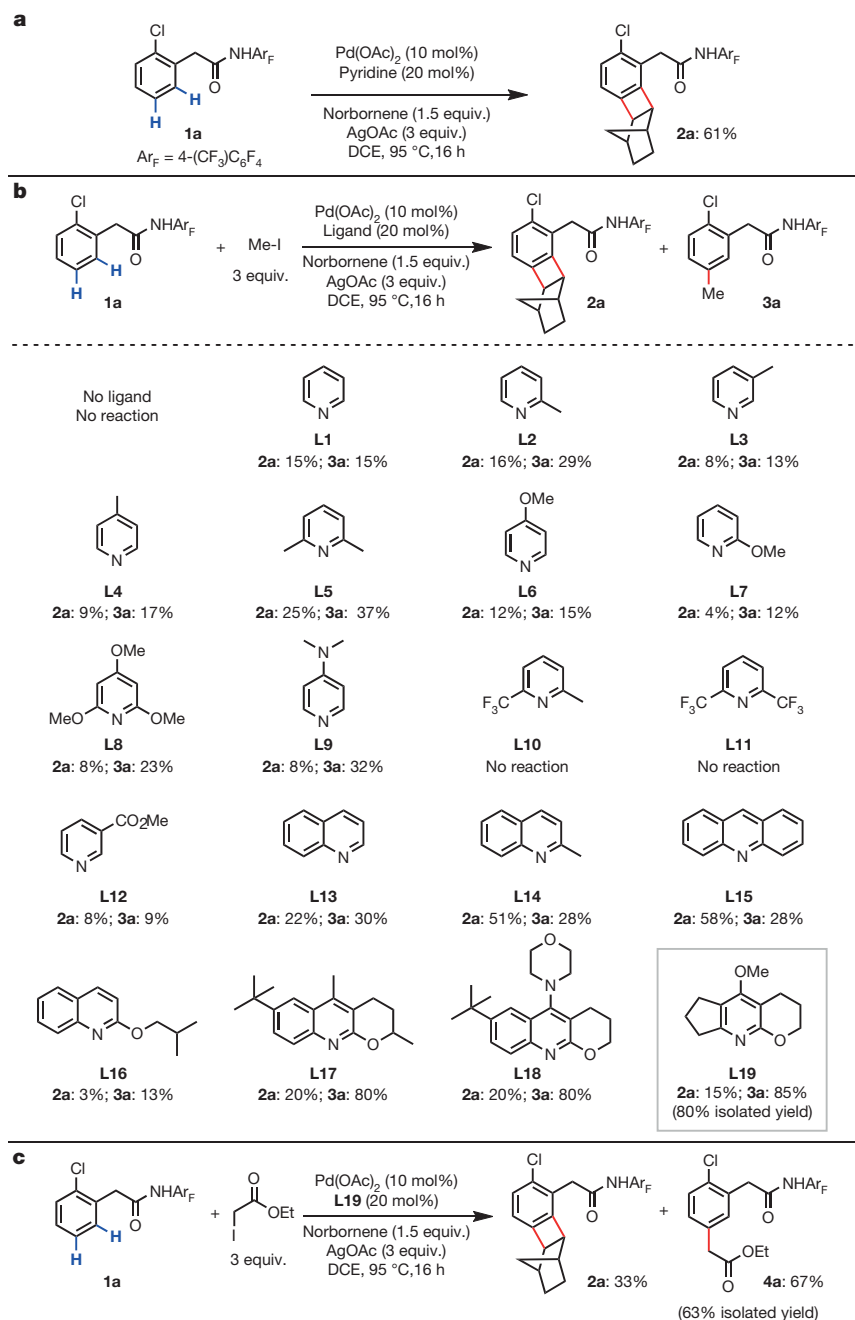


Figure 2 | Discovery of a ligand that enables *meta*-C–H alkylation using norbornene as a transient mediator. **a**, Pyridine promotes the Pd-catalysed reaction of phenylacetic amide **1a** with norbornene to afford benzocyclobutene via a Catellani pathway. DCE, 1,2-dichloroethane. **b**, A modified pyridine-type

ligand enables selective *meta*-C–H methylation. Yields were determined by ¹H NMR analysis of the crude reaction mixture. **c**, *meta*-C–H alkylation with ethyl iodoacetate.

We began our investigation with a phenylacetic-acid-derived *N*-2,3,5,6-tetrafluoro-4-trifluoromethylphenyl amide substrate. This auxiliary has proved to be effective in the *ortho*-C–H activation of $C(sp^2)$ -H bonds by Pd(II) catalysts¹². To probe the feasibility of each step in the proposed catalytic cycle, we first treated substrate **1a** with Pd(OAc)₂ (OAc = acetate) and norbornene under various conditions in an attempt to identify conditions for 1,2-migratory insertion without success (see Supplementary Information). Although *ortho*-C–H activation of **1a** does not require a ligand²⁶, our recent finding that 9-methylacridine promotes directed *ortho*-C–H activation²⁷ prompted us to search for pyridine-type ligands that would facilitate the 1,2-migratory insertion. We found that the addition of pyridine as a ligand enabled the 1,2-migratory insertion of norbornene into the putative *ortho*-palladacycle intermediate. Subsequent *meta*-C–H palladation and reductive elimination presumably then took place to give benzocyclobutene **2a** in 61% yield (Fig. 2a). Formation of **2a** was not observed in the absence of pyridine ligand.

Having validated the first key steps in the proposed catalytic cycle (*ortho*-C–H activation, 1,2-migratory insertion, and *meta*-C–H activation), we next sought to identify conditions to prevent benzocyclobutene formation and promote *meta*-functionalization with an electrophile. In the presence of methyl iodide, the desired *meta*-methylated product **3a** was obtained in 15% yield, and the yield of **2a** decreased to 15% (Fig. 2b, **L1**). Further screening of simple pyridine ligands revealed that alkyl substitution at the 2-position improved the yield of *meta*-methylation (**L2** and **L5**, 29% and 37% yield, respectively). We further found that pyridine ligands with electron-donating substituents gave improved product selectivity (**L6–L9**), whereas electron-deficient pyridine ligands provided both poor selectivity and reactivity (**L10–L12**). Prompted by our recent progress in developing quinoline-based ligands for C–H activation^{28,29}, we subsequently examined this class of ligands. Although

the use of 2-methylquinoline (**L14**) and acridine (**L15**) as ligands led to increased reactivity, formation of benzocyclobutene **2a** was favoured. Because electron-donating substituents enhance selectivity for the *meta*-methylation pathway, we tested several quinoline ligands containing alkoxy groups at the 2-position. Although 2-(isobutoxy)quinoline (**L16**) was an ineffective ligand, the use of ligand **L17**, containing a 2-alkoxy group with a constrained conformation in a fused ring, led to full conversion, affording **3a** in 80% yield. Installation of a highly electron-donating amine substituent on a similar ligand **L18** did not improve the selectivity. Ligand **L19**, containing a fused cyclopentane ring, further improved the selectivity for *meta*-methylation, providing **3a** in 85% yield. Selective *meta*-alkylation with ethyl iodoacetate was also achieved under optimized conditions to give the desired product **4a** in 67% yield (Fig. 2c).

The scope of this norbornene-mediated *meta*-alkylation was investigated, using methyl iodide (**3a–3x**) and ethyl iodoacetate (**4a–4x**) as coupling partners (Fig. 3). Electron-donating and electron-withdrawing *ortho*-substituents were tolerated in this transformation, providing moderate to excellent yields (**3a–3h** and **4a–4h**). An *ortho*-bromide substituent was also compatible with this protocol (**3d** and **4d**). Alkylation of 1-naphthylacetic amide **1i** occurred exclusively at the 3-position, rather than at the 8-position (**3i** and **4i**). When an unsubstituted amide substrate **1j** was used, methylation occurred at both *meta* positions to provide dimethylated product **3j** in 80% yield. Coupling of ethyl iodoacetate with this substrate, however, mainly produced the benzocyclobutene by-product; the yield of the desired dialkylated product **4j** was less than 10%. This result demonstrates the subtle reactivity differences between these two alkylating reagents. Substrates bearing electron-donating or electron-withdrawing *meta*-substituents were also reactive in this protocol (**3k–3n** and **4k–4n**). Methylation of the more sterically

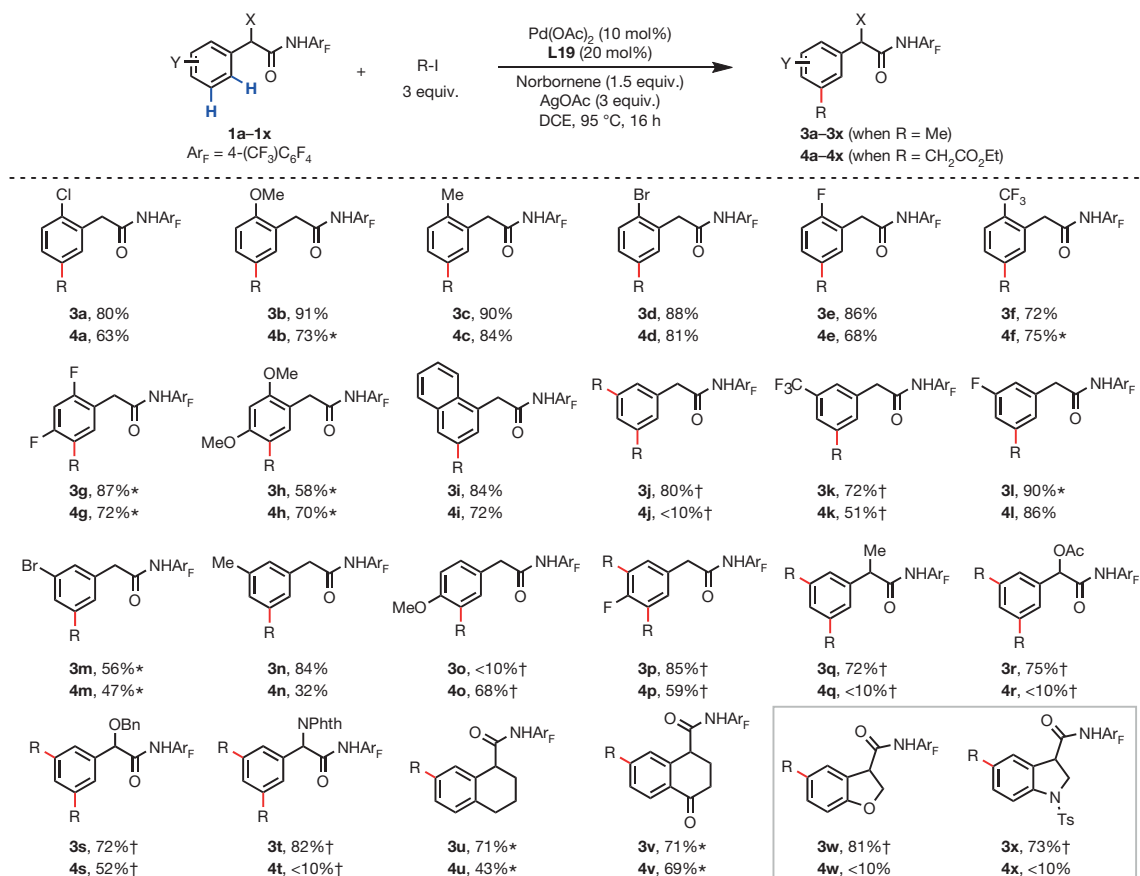


Figure 3 | *meta*-C–H alkylation of phenylacetic amides. R-I, methyl iodide or ethyl iodoacetate; X, Y, generic substituents; Ac, acetyl; Bn, benzyl; Phth, phthaloyl; Ts, 4-toluenesulfonyl. The values under each structure indicate isolated yields. For **1j** and **1p–1t**, dialkylation is observed; reaction conditions:

20 mol% Pd(OAc)₂, 40 mol% **L19**, 5 equiv. R-I, 5 equiv. AgOAc, 3 equiv. norbornene. *, 15 mol% Pd(OAc)₂, 30 mol% **L19**. †, 20 mol% Pd(OAc)₂, 40 mol% **L19**.

hindered *p*-MeO-substituted substrate **1o** was sluggish and poor yielding, providing **3o** in less than 10% yield. Mono-alkylation of **1o** with ethyl iodoacetate afforded **4o** in 68% yield. A substrate containing the less sterically bulky *p*-F-group reacted well with both coupling partners to afford the dialkylated products **3p** and **4p** in 85% and 59% yields, respectively. The methylation reaction proceeded effectively with a broad range of α -substituents, including methyl, acetoxyl, benzyloxy and phthaloyl-protected amino groups to give the dimethylated products in good yields (**3q–3t**). Thus, this transformation enabled the direct *meta*-functionalization of key biologically active building blocks, such as mandelic acid (**3r** and **3s**) and phenylglycine (**3t**). In contrast, the ethoxycarbonylmethylation reaction was effective only for **1s**, containing a benzyloxy group at the benzylic position (**4s**). Both alkylation reactions were compatible with tetrahydronaphthalene (**1u**) and tetralone (**1v**) substrates, despite the amide directing groups being out of plane

with the target C(aryl)–H bonds. *Meta*-methylation was effective for the heterocyclic dihydrobenzofuran-containing and indoline-containing substrates **1w** and **1x**. However, ethoxycarbonylmethylation of these heterocycles provided only trace product.

We subsequently examined the generality of this *meta*-C–H functionalization strategy by investigating additional organohalide coupling partners (Fig. 4). Ethylation with ethyl iodide under standard conditions led to exclusive formation of the undesired benzocyclobutene by-product. Increasing the loading of ethyl iodide (6 equiv.) afforded *meta*-ethylated product **5** in 21% yield, with the benzocyclobutene still being formed as the major product. *Meta*-alkylation with benzyl bromide, a more reactive electrophile, proceeded smoothly to afford **6a** in 66% yield. Both electron-donating and electron-withdrawing substituents on the *ortho*, *meta* and *para* positions of the benzyl bromide coupling partner were tolerated, providing *meta*-benzylated products **6b–6g** in 60–75% yield.

Finally, we examined whether this approach for *meta*-C–H functionalization could be adapted to achieve *meta*-C–H arylation, thereby providing a method for the synthesis of biaryl compounds. Arylation with methyl 2-iodobenzoate in *t*-butyl methyl ether (TBME) provided the *meta*-arylated product **7a** in 64% yield. The use of methyl 3-iodobenzoate or methyl 4-iodobenzoate, however, resulted in the formation of the benzocyclobutene adduct as the major product. These experiments suggest that the presence of an *ortho*-coordinating group is required to promote oxidative addition of the aryl halide, which is consistent with previous studies^{19,30}. We therefore examined other substituted 2-iodobenzoates and were able to obtain the desired products in moderate yields (**7b–7d**). Similarly, 2'-iodoacetophenone was a competent coupling partner, affording **7e** in 48% yield. Arylation with highly reactive, electron-deficient 3,5-bis(trifluoromethyl)iodobenzene was also effective, providing the *meta*-arylated product **7f** in 76% yield. The successful *meta*-arylation of substrates containing dihydrobenzofuran (**8**, 82% yield) and indoline (**9**, 69% yield) groups with methyl 2-iodobenzoate is especially noteworthy in terms of demonstrating late-stage diversification of heterocycles with possible applications in medicinal chemistry. We expect that further development of the ancillary pyridine ligand will broaden the scope of alkyl and aryl halides in this reaction, ultimately leading to a broadly useful method for *meta*-C–H functionalization using common *ortho*-directing groups and norbornene as a transient mediator.

Received 26 October; accepted 31 December 2014.

Published online 9 March 2015.

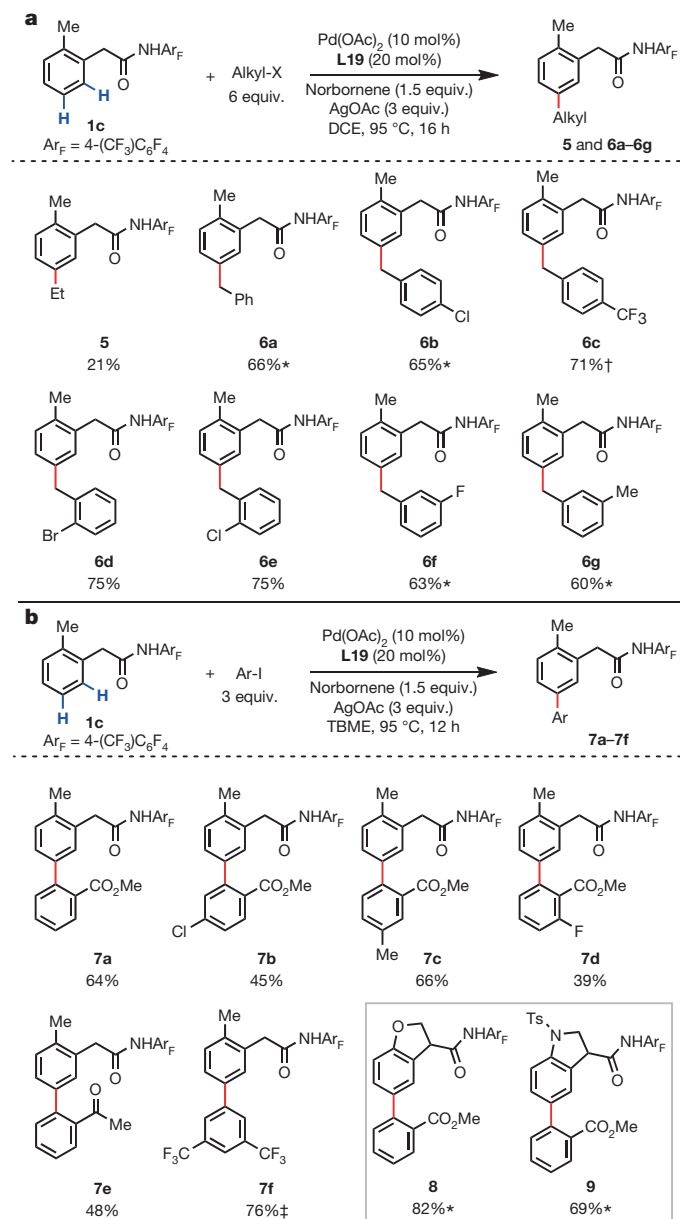


Figure 4 | Scope of organohalide coupling partners. **a**, *meta*-C–H alkylation with alkyl halides. Alkyl-X, iodoethane or benzyl bromides. **b**, *meta*-C–H arylation with aryl iodides. TBME, *t*-butyl methyl ether. The values under each structure indicate isolated yields. *, 15 mol% Pd(OAc)₂, 30 mol% L19. †, Dichloromethane as solvent. ‡, 6 equiv. Ar-I; AgOPiv (OPiv = pivalate) used instead of AgOAc.

- Leow, D., Li, G., Mei, T.-S. & Yu, J.-Q. Activation of remote *meta*-C–H bonds assisted by an end-on template. *Nature* **486**, 518–522 (2012).
- Tang, R.-Y., Li, G. & Yu, J.-Q. Conformation-induced remote *meta*-C–H activation of amines. *Nature* **507**, 215–220 (2014).
- Lee, S., Lee, H. & Tan, K. L. *Meta*-selective C–H functionalization using a nitrile-based directing group and cleavable Si-tether. *J. Am. Chem. Soc.* **135**, 18778–18781 (2013).
- Wan, L., Dashtbaravardeh, N., Li, G. & Yu, J.-Q. Cross-coupling of remote *meta*-C–H bonds directed by a U-shaped template. *J. Am. Chem. Soc.* **135**, 18056–18059 (2013).
- Yang, G. *et al.* Pd(II)-catalyzed *meta*-C–H olefination, arylation, and acetoxylation of indolines using a U-shaped template. *J. Am. Chem. Soc.* **136**, 10807–10813 (2014).
- Brown, J. M. Selectivity and mechanism in catalytic asymmetric synthesis. *Chem. Soc. Rev.* **22**, 25–41 (1993).
- Johnson, R. A. & Sharpless, K. B. in *Catalytic Asymmetric Synthesis* 2nd edn (ed. Ojima, I.) 231–280 (Wiley, 2005).
- Li, Z. & Yamamoto, H. Hydroxamic acids in asymmetric synthesis. *Acc. Chem. Res.* **46**, 506–518 (2013).
- Hartung, C. G. & Snieckus, V. in *Modern Arene Chemistry* (ed. Astruc, D.) 330–367 (Wiley-VCH, 2004).
- Lyons, T. W. & Sanford, M. S. Palladium-catalyzed ligand-directed C–H functionalization reactions. *Chem. Rev.* **110**, 1147–1169 (2010).
- Daugulis, O., Do, H.-Q. & Shabashov, D. Palladium- and copper-catalyzed arylation of carbon–hydrogen bonds. *Acc. Chem. Res.* **42**, 1074–1086 (2009).
- Engle, K. M., Mei, T.-S., Wasa, M. & Yu, J.-Q. Weak coordination as a powerful means for developing broadly useful C–H functionalization reactions. *Acc. Chem. Res.* **45**, 788–802 (2012).
- Colby, D. A., Bergman, R. G. & Ellman, J. A. Rhodium-catalyzed C–C bond formation via heteroatom-directed C–H bond activation. *Chem. Rev.* **110**, 624–655 (2010).

14. Wencel-Delord, J., Dröge, T., Liu, F. & Glorius, F. Towards mild metal-catalyzed C–H bond activation. *Chem. Soc. Rev.* **40**, 4740–4761 (2011).
15. Saidi, O. *et al.* Ruthenium-catalyzed *meta*-sulfonation of 2-phenylpyridines. *J. Am. Chem. Soc.* **133**, 19298–19301 (2011).
16. Duong, H. A., Gilligan, R. E., Cooke, M. L., Phipps, R. J. & Gaunt, M. J. Copper(II)-catalyzed *meta*-selective direct arylation of α -aryl carbonyl compounds. *Angew. Chem. Int. Ed.* **50**, 463–466 (2011).
17. Hofmann, N. & Ackermann, L. *meta*-Selective C–H bond alkylation with secondary alkyl halides. *J. Am. Chem. Soc.* **135**, 5877–5884 (2013).
18. Catellani, M., Frignani, F. & Ragoni, A. A complex catalytic cycle leading to a regioselective synthesis of *o,o'*-disubstituted vinylarenes. *Angew. Chem. Int. Edn Engl.* **36**, 119–122 (1997).
19. Faccini, F., Motti, E. & Catellani, M. A new reaction sequence involving palladium-catalyzed unsymmetrical aryl coupling. *J. Am. Chem. Soc.* **126**, 78–79 (2004).
20. Martins, A., Mariampillai, B. & Lautens, M. Synthesis in the key of Catellani: norbornene-mediated *ortho* C–H functionalization. *Top. Curr. Chem.* **292**, 1–33 (2010).
21. Cárdenas, D. J., Martín-Matute, B. & Echavarren, A. M. Aryl transfer between Pd(II) centers or Pd(IV) intermediates in Pd-catalyzed domino reactions. *J. Am. Chem. Soc.* **128**, 5033–5040 (2006).
22. Dong, Z. & Dong, G. *Ortho* vs *ipso*: site-selective Pd and norbornene-catalyzed arene C–H amination using aryl halides. *J. Am. Chem. Soc.* **135**, 18350–18353 (2013).
23. Jiao, L. & Bach, T. Palladium-catalyzed direct 2-alkylation of indoles by norbornene-mediated regioselective cascade C–H activation. *J. Am. Chem. Soc.* **133**, 12990–12993 (2011).
24. Jiao, L., Herdtweck, E. & Bach, T. Pd(II)-catalyzed regioselective 2-alkylation of indoles via a norbornene-mediated C–H activation: mechanism and applications. *J. Am. Chem. Soc.* **134**, 14563–14572 (2012).
25. Catellani, M. & Ferioli, L. An improved synthesis of 1,4-*cis,exo*-hexa- or tetrahydromethano- or -ethanobiphenylene derivatives catalyzed by palladium complexes. *Synthesis* 769–772 (1996).
26. Wang, X.-C. *et al.* Pd(II)-catalyzed C–H iodination using molecular I₂ as the sole oxidant. *J. Am. Chem. Soc.* **135**, 10326–10329 (2013).
27. Zhu, R.-Y., He, J., Wang, X.-C. & Yu, J.-Q. Ligand-promoted alkylation of C(sp³)–H and C(sp²)–H bonds. *J. Am. Chem. Soc.* **136**, 13194–13197 (2014).
28. He, J. *et al.* Ligand-controlled C(sp³)–H arylation and olefination in synthesis of unnatural chiral α -amino acids. *Science* **343**, 1216–1220 (2014).
29. Li, S., Chen, G., Feng, C.-G., Gong, W. & Yu, J.-Q. Ligand-enabled γ -C–H olefination and carbonylation: construction of β -quaternary carbon centers. *J. Am. Chem. Soc.* **136**, 5267–5270 (2014).
30. Martins, A., Candito, D. A. & Lautens, M. Palladium-catalyzed reductive *ortho*-arylation: evidence for the decomposition of 1,2-dimethoxyethane and subsequent arylpalladium(II) reduction. *Org. Lett.* **12**, 5186–5188 (2010).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank J. Spangler and A. Horns for constructive suggestions. We thank the Scripps Research Institute and the National Institutes of Health (NIGMS, 1R01 GM102265) for financial support. L.-Z.F. is a visiting scholar from School of Pharmacy, Xinxiang Medical University and is sponsored by the China Scholarship Council.

Author Contributions X.-C.W. and R.-Y.Z. developed the reactions. W.G. and S.L. synthesized the ligands. L.-Z.F. examined the substrate scope. K.M.E. performed preliminary studies. J.-Q.Y. conceived this concept and prepared the manuscript with feedback from X.-C.W.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.-Q.Y. (yu200@scripps.edu).

Observational determination of surface radiative forcing by CO₂ from 2000 to 2010

D. R. Feldman¹, W. D. Collins^{1,2}, P. J. Gero³, M. S. Torn^{1,4}, E. J. Mlawer⁵ & T. R. Shippert⁶

The climatic impact of CO₂ and other greenhouse gases is usually quantified in terms of radiative forcing¹, calculated as the difference between estimates of the Earth's radiation field from pre-industrial and present-day concentrations of these gases. Radiative transfer models calculate that the increase in CO₂ since 1750 corresponds to a global annual-mean radiative forcing at the tropopause of $1.82 \pm 0.19 \text{ W m}^{-2}$ (ref. 2). However, despite widespread scientific discussion and modelling of the climate impacts of well-mixed greenhouse gases, there is little direct observational evidence of the radiative impact of increasing atmospheric CO₂. Here we present observationally based evidence of clear-sky CO₂ surface radiative forcing that is directly attributable to the increase, between 2000 and 2010, of 22 parts per million atmospheric CO₂. The time series of this forcing at the two locations—the Southern Great Plains and the North Slope of Alaska—are derived from Atmospheric Emitted Radiance Interferometer spectra³ together with ancillary measurements and thoroughly corroborated radiative transfer calculations⁴. The time series both show statistically significant trends of 0.2 W m^{-2} per decade (with respective uncertainties of $\pm 0.06 \text{ W m}^{-2}$ per decade and $\pm 0.07 \text{ W m}^{-2}$ per decade) and have seasonal ranges of $0.1\text{--}0.2 \text{ W m}^{-2}$. This is approximately ten per cent of the trend in downwelling longwave radiation^{5–7}. These results confirm theoretical predictions of the atmospheric greenhouse effect due to anthropogenic emissions, and provide empirical evidence of how rising CO₂ levels, mediated by temporal variations due to photosynthesis and respiration, are affecting the surface energy balance.

Even though Northern Hemisphere atmospheric CO₂ mixing ratios have recently exceeded 400 parts per million (ppm), few investigations have directly explored the effect of changes in well-mixed greenhouse gases on spectral and broadband radiation fluxes. In fact, doing so has proved difficult. Using data from two satellite instruments launched 26 years apart, Harries *et al.*⁸ attributed systematically decreased emission in CO₂ and CH₄ spectral bands to increased opacity from, and rising concentrations in, these well-mixed greenhouse gases. However, this effort was complicated by uncertainties in instrument performance, short measurement records from each instrument, and cloud contamination⁹.

In principle, CO₂ forcing can be predicted from knowledge of the atmospheric state assuming exact spectroscopy and accurate radiative transfer. Forcing can then be estimated using radiative transfer calculations with atmospheric temperature, the concentrations of radiatively active constituents including water vapour, O₃, CH₄, N₂O, and less prominent well-mixed greenhouse gases, and changes in CO₂. However, experimental validation of this forcing is needed outside the laboratory because CO₂ spectroscopy is an area of active research^{10–13}. Furthermore, the fast radiative-transfer algorithms that drive regional and global climate models approximate spectroscopic absorption line-by-line calculations with errors of about 0.6 W m^{-2} (ref. 14), an amount comparable to the forcing by anthropogenic CH₄ and N₂O.

Surface forcing represents a complementary, underutilized resource with which to quantify the effects of rising CO₂ concentrations on

downwelling longwave radiation. This quantity is distinct from stratosphere-adjusted radiative forcing at the tropopause, but both are fundamental measures of energy imbalance caused by well-mixed greenhouse gases¹⁵. The former is less than, but proportional to, the latter owing to tropospheric adjustments of sensible and latent heat¹⁶, and is a useful metric for localized aspects of climate response². We focus here on clear-sky flux changes because models predict most of the CO₂ surface forcing to occur under clear-sky conditions¹⁵.

Specialized atmospheric observations at experimental sites in the mid-latitude continental Southern Great Plains (SGP) and the Arctic marine North Slope of Alaska (NSA) sites by the US Department of Energy Atmospheric Radiation Measurement (ARM) programme¹⁷ produce the integrated data sets required for an independent diagnosis of the surface radiative effects of CO₂. We used spectroscopic measurements from the Atmospheric Emitted Radiance Interferometer (AERI) instrument and atmospheric state data at these two sites to test whether the impact of rising CO₂ on downwelling longwave radiation can be rigorously detected. By basing this analysis on sets of independent measurements at high temporal frequency over long duration, we empirically established how anthropogenic emissions, mediated by variations in photosynthetic activity and respiration, are altering the Earth's surface energy balance.

However, AERI spectral measurements and trends are sensitive to many different components of the atmospheric state. To interpret these measurements and attribute specific signals to rising CO₂ requires an accurate radiative-transfer model that reproduces these spectra on the basis of an independent assessment of the state of the atmosphere. The model must capture instantaneous signals and long-term trends in the spectra to determine the effects of CO₂ on diurnal to decadal timescales.

We used the Line-by-Line Radiative Transfer Model (LBLRTM)¹⁸, which is continuously compared against other line-by-line models⁴ and observations¹⁹. A sample clear-sky measured AERI spectrum is shown in Fig. 1a. Figure 1b shows residual spectra produced from the measurement ('obs'), minus spectra calculated ('calc') using (1) CO₂ concentrations from CarbonTracker 2011 (CT2011)²⁰, which is a greenhouse gas assimilation system based on measurements and modelled emission and transport; (2) methane (CH₄) profiles from CarbonTracker-CH₄ (ref. 21); (3) ozone (O₃) profiles from NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA)²²; and (4) temperature and water-vapour profiles from radiosondes (see Methods). The measured spectrum in Fig. 1a shows Planck function behaviour near the centre of the fundamental (ν_2) CO₂ band and exhibits a departure from a Planck curve in the P- and R-branches of this feature, indicating that the emission in these branches is sub-saturated and could increase with increasing CO₂. Water-vapour features, continuum emission, and O₃ emission are seen in the infrared window between 800 cm^{-1} and $1,200 \text{ cm}^{-1}$, and lesser features from CH₄ are seen around $1,300 \text{ cm}^{-1}$. Calculated transmission and the change in transmission with a 22 ppm CO₂ increase are also shown, indicating that weak vibration-rotation features in the far wings of the

¹Lawrence Berkeley National Laboratory, Earth Sciences Division, 1 Cyclotron Road, MS 74R-316C, Berkeley, California 94720, USA. ²University of California-Berkeley, Department of Earth and Planetary Science, 307 McCone Hall, MC 4767, Berkeley, California 94720, USA. ³University of Wisconsin-Madison, Space Science and Engineering Center, 1225 W. Dayton Street, Madison, Wisconsin 53706, USA. ⁴University of California-Berkeley, Energy and Resources Group, Berkeley, 310 Barrows Hall, MC 3050, California 94720, USA. ⁵Atmospheric and Environmental Research, Inc., 131 Hartwell Avenue, Lexington, Massachusetts 02141, USA. ⁶Pacific Northwest National Laboratory, Fundamental and Computational Sciences, 902 Battelle Boulevard, Richland, Washington 99354, USA.

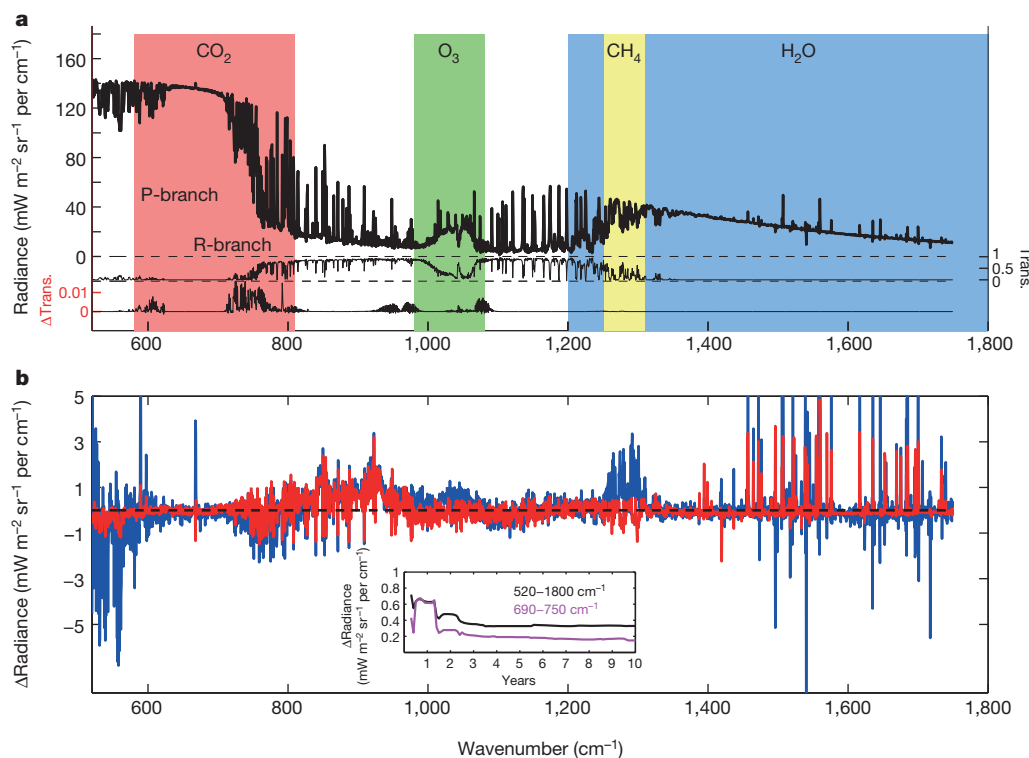


Figure 1 | AERI spectrum and residual features. **a**, Sample clear-sky AERI (channel 1) spectrum measured at SGP on 14 March 2001 2330Z, transmission calculation (Trans.) from LBLRTM, and the difference in transmission (Δ Trans.) calculated for a 22 ppm change in column-averaged CO_2 (370–392 ppm). **b**, The

instantaneous spectral residual (obs – calc) (blue) and a spectral residual from observations for March 2001 (red). The inset indicates the running average of spectral residual rms of AERI Channel 01 (520–1,800 cm^{-1}) and the CO_2 R-branch (690–750 cm^{-1}).

fundamental and in the infrared window dominate surface radiative forcing from rising CO_2 .

The agreement between a single measured spectrum and the LBLRTM calculation (that is, the residual) is generally within $1.5 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ per cm}^{-1}$, with notable exceptions. These include the region below 550 cm^{-1} where instrumental thermal noise is present, the centre of the ν_2 CO_2 absorption feature, which is sensitive to the temperature within the instrument container, and sporadic features between $1,400 \text{ cm}^{-1}$ and $1,800 \text{ cm}^{-1}$ that are due to opaque water vapour lines. Nevertheless, over 90% of the residual is within a 3σ envelope established by the noise-effective radiance of $0.2 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ per cm}^{-1}$. Noise in the residual spectrum does not preclude long-term spectral analysis²³ because it has no long-term bias. As highlighted in the inset of Fig. 1 the spectral root mean square (rms) of the residual decreases with the length of time over which the spectra are averaged, with asymptotes at $0.3 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ per cm}^{-1}$ and $0.1 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ per cm}^{-1}$ across the AERI band-pass and in the CO_2 R-branch, respectively.

Over the length of the observation period (2000–2010), the modelled spectra at both SGP and NSA are dominated by trends associated with the temperature and humidity structure of the atmosphere rather than the smaller signal from CO_2 . The seasonal and annual trends in calculated clear-sky spectra at SGP (Fig. 2a) and NSA (Fig. 2d) are dominated by changes in the atmospheric thermodynamic state and are of opposite sign depending on the season. These signals arise from seasonally dependent clear-sky trends in temperature profiles and water vapour concentrations, as determined by radiosondes (see Methods) and must be taken into account to determine the forcing from CO_2 . We therefore construct counterfactual spectra (such spectra are produced from models that keep the CO_2 concentration fixed) to simulate spectra with time-invariant CO_2 , whereby we use temperature and water-vapour estimates from concurrent radiosondes to remove the thermodynamically derived radiometric signals from AERI spectra and isolate the signature of CO_2 . Since most CO_2 surface forcing occurs in the absence of clouds¹⁶,

we focus on clear-sky conditions, identified using the Active Remote-Sensing of Clouds (ARSCL) Value-Added Product²⁴ data set.

Differences between counterfactual spectra and coincidental AERI measurements show structure in the major CO_2 absorption features as shown in Fig. 2b and e, at an order of magnitude greater than the long-term residual rms in Fig. 1b. Also shown are those spectral features for which the trend is non-zero at the 3σ level²⁵. These panels show the unmistakable spectral fingerprint of CO_2 . The trends in forcing are significantly ($P < 0.003$) different from zero only in the P- and R- branches of the ν_2 CO_2 band.

We can exclude alternative explanations for the change in these measurements, such as instrument calibration or the temperature, water vapour, or condensate structure of the atmosphere because they would produce significant ($P < 0.003$) trends in other spectral regions outside the CO_2 absorption bands—see Fig. 2b and e. Moreover, the spectral forcing from CO_2 is a strong function of changes in the CO_2 column concentration, and nonlinear interactions between temperature and water vapour were weak, as indicated by the lack of statistically significant differences in the seasonal and annual spectral trends in the CO_2 P- and R-branches. Therefore, the atmospheric structure of temperature and water vapour does not strongly affect CO_2 surface forcing, which is consistent with the findings of others^{26,27}.

Finally, when we compare counterfactual spectra with radiative transfer calculations where CO_2 is derived from CT2011, as shown in Fig. 2c and f, we again find spectral structure only in the CO_2 absorption features, thereby confirming that the model is a reliable tool for determining the surface forcing from CO_2 . The utility of the counterfactual approach is further demonstrated by the probability distribution function of the difference in the rms spectral residual of AERI and LBLRTM with time-varying and fixed CO_2 concentrations for 2010 (Fig. 3). The two probability distribution functions differ substantially (two-sided t -test, $P < 0.00001$). The distribution of the mean residuals also differ significantly (two-sided t -test, $P < 0.00001$).

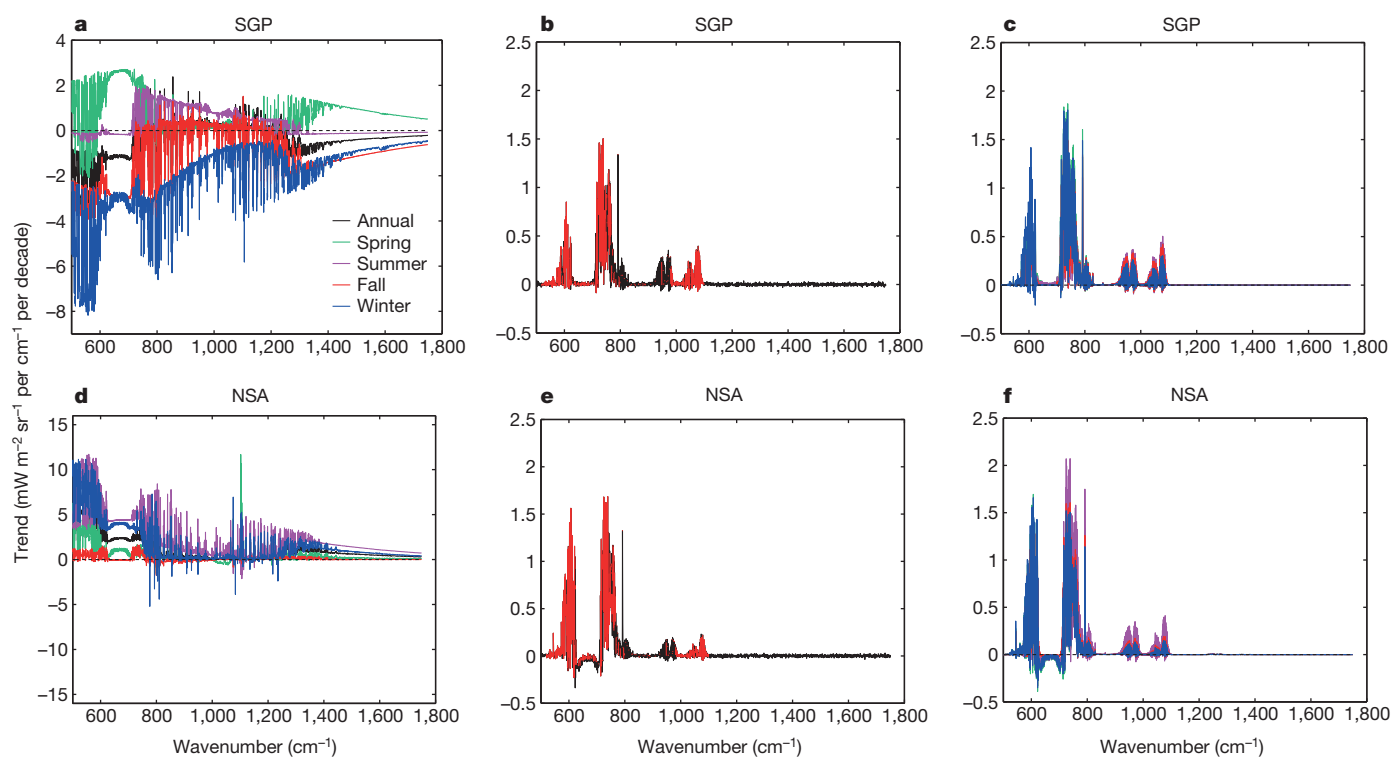


Figure 2 | Measured and modelled spectral trends for 2000–2010.

a, Calculated (simulated) SGP AERI clear-sky seasonal and annual spectral radiance trends for 2000–2010 with temperature, water vapour, and CO₂ changes. **b**, SGP AERI annually averaged clear-sky spectral residual trends, where the residual is the difference between the measurement and a

radiative-transfer calculation with CO₂ fixed at 370 ppm (in black). Red overlay indicates channel trends 3σ different from zero. **c**, As for **b** but showing residual trends from differencing calculations based on CT2011 CO₂ concentrations and CO₂ fixed at 370 ppm. **d**, As for **a** but for NSA. **e**, As for **b** but for NSA. **f**, As for **c** but for NSA.

The time series of CO₂ surface forcing, derived from differencing AERI measurements and counterfactual calculations at the SGP (Fig. 4a) and spectrally integrating and converting to flux, shows clear and increasing trends in radiative surface forcing and seasonal variability. The least-squares trend in the long-term forcing is $0.2 \pm 0.06 \text{ W m}^{-2}$ per decade and differs significantly ($P < 0.003$) from zero. The seasonal amplitude of the forcing is $0.1\text{--}0.2 \text{ W m}^{-2}$, closely tracking the independently assessed pattern in the average CO₂ concentration in the

lowest 2 km of the atmosphere¹⁹. The variation in the power spectral density function of surface forcing with frequency (Fig. 4b) shows the largest peak associated with springtime photosynthesis and autumn respiration.

The time-series for NSA (Fig. 4c) also has a pronounced seasonal cycle and secular trend. The range of CO₂ surface forcing is similar to that of SGP, although the higher frequency variability is much less prominent. In addition, the time series of surface forcing at NSA shows increasing variability in the latter part of the 11-year analysis record. This variability results from increased numbers of samples and fewer outages (when not all of the necessary data streams are available to derive the CO₂ forcing) at NSA from 2004 onwards. Nevertheless, the least-squares trend in CO₂ surface forcing at NSA is also $0.2 \pm 0.07 \text{ W m}^{-2}$ per decade, with a seasonal range of 0.1 W m^{-2} and differs significantly from zero ($P < 0.02$).

Increasing atmospheric CO₂ concentrations between 2000 and 2010 have led to increases in clear-sky surface radiative forcing of over 0.2 W m^{-2} at mid- and high-latitudes. Fossil fuel emissions and fires contributed substantially to the observed increase²⁰. The climate perturbation from this surface forcing will be larger than the observed effect, since it has been found that the water-vapour feedback enhances greenhouse gas forcing at the surface by a factor of three²⁸ and will increase, largely owing to thermodynamic constraints²⁹. The evolving roles of atmospheric constituents, including water vapour and CO₂ (ref. 30), in their radiative contributions to the surface energy balance can be tracked with surface spectroscopic measurements from stand-alone (or networks of) AERI instruments. If CO₂ concentrations continue to increase at the current mean annual rate of 2.1 ppm per year, these spectroscopic measurements will continue to provide robust evidence of radiative perturbations to the Earth's surface energy budget due to anthropogenic climate change, but mediated by annual variations in photosynthetic activity. These perturbations will probably influence other energy fluxes and key properties of the Earth's surface and should be explored further.

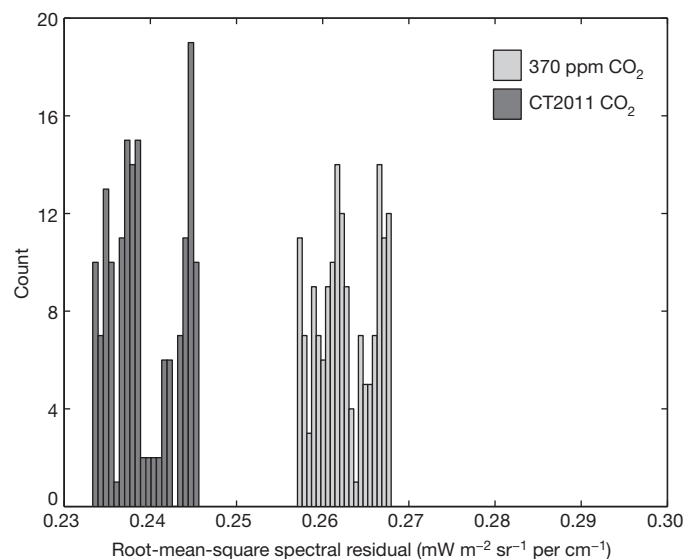


Figure 3 | Distributions of residual rms values in 2010. Histograms of rms spectral residuals ($520\text{--}1,800 \text{ cm}^{-1}$) at SGP with time-varying CO₂ (dark grey) and a uniform 370 ppm CO₂ profile (light grey). Each count corresponds to a separate spectrum measured by AERI in 2000 or 2010.

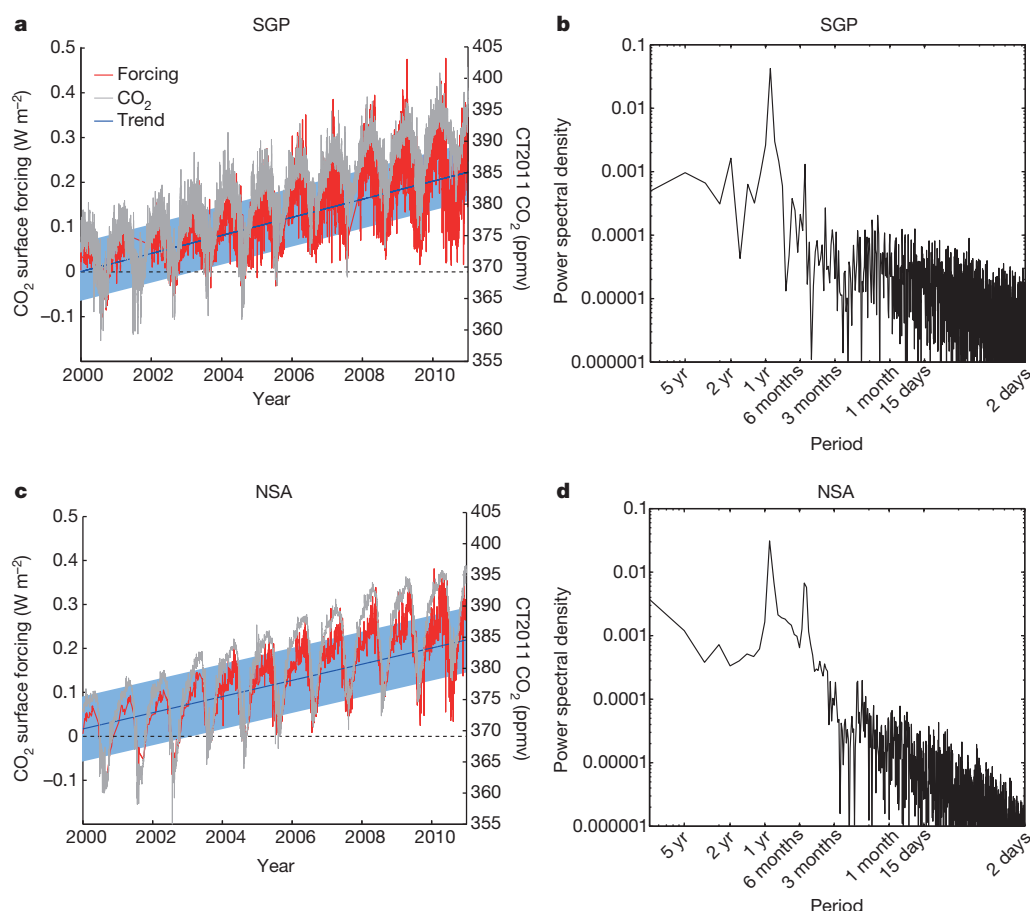


Figure 4 | Time-series of surface forcing. **a**, Time series of observed spectrally integrated (520–1,800 cm⁻¹) CO₂ surface radiative forcing at SGP (in red) with overlaid CT2011 estimate of CO₂ concentration from the surface to an

altitude of 2 km (grey), and a least-squares trend of the forcing and its uncertainty (blue). **b**, Power spectral density of observed CO₂ surface radiative forcing at SGP. **c**, As for **a** but for the NSA site. **d**, As for **b** but for the NSA site.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 June 2014; accepted 15 January 2015.

Published online 25 February 2015.

- Ramaswamy, V. *et al.* in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (eds Houghton, J. T. *et al.*) 349–416 (Cambridge Univ. Press, 2001).
- Myhre, G. *et al.* Anthropogenic and natural radiative forcing. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. *et al.*) 661 (Cambridge Univ. Press, 2013).
- Knuteson, R. O. *et al.* Atmospheric emitted radiance interferometer. Part I: Instrument design. *J. Atmos. Ocean. Technol.* **21**, 1763–1776 (2004).
- Oreopoulos, L. *et al.* The continual intercomparison of radiation codes: results from phase I. *J. Geophys. Res.* **117**, D06118 (2012).
- Prata, F. The climatological record of clear-sky longwave radiation at the Earth's surface: evidence for water vapour feedback? *Int. J. Remote Sens.* **29**, 5247–5263 (2008).
- Wild, M., Grieser, J. & Schär, C. Combined surface solar brightening and increasing greenhouse effect support recent intensification of the global land-based hydrological cycle. *Geophys. Res. Lett.* **35**, L17706 (2008).
- Wang, K. & Liang, S. Global atmospheric downward longwave radiation over land surface under all-sky conditions from 1973 to 2008. *J. Geophys. Res.* **114**, D19101 (2009).
- Harries, J. E., Brindley, H. E., Sagoo, P. J. & Bantges, R. J. Increases in greenhouse forcing inferred from the outgoing longwave radiation spectra of the Earth in 1970 and 1997. *Nature* **410**, 355–357 (2001).
- Jiang, Y., Aumann, H. H., Wingyee-Lau, M. & Yung, Y. L. Climate change sensitivity evaluation from AIRS and IRIS measurements. *Proc. SPIE* **8153**, XVI, <http://dx.doi.org/10.1117/12.892817> (2011).
- Rothman, L. S. *et al.* The HITRAN2012 molecular spectroscopic database. *J. Quant. Spectrosc. Radiat.* **130**, 4–50 (2013).
- Kochel, J.-M., Hartmann, J.-M., Camy-Peyret, C., Rodrigues, R. & Payan, S. Influence of line mixing on absorption by CO₂ Q branches in atmospheric balloon-borne spectra near 13 μm. *J. Geophys. Res.* **102** (D11), 12891–12899 (1997).
- Niro, F., Jucks, K. & Hartmann, J.-M. Spectra calculations in central and wing regions of CO₂ IR bands. IV: Software and database for the computation of atmospheric spectra. *J. Quant. Spectrosc. Radiat.* **95**, 469–481 (2005).
- Alvarado, M. J. *et al.* Performance of the Line-By-Line Radiative Transfer Model (LBLRTM) for temperature, water vapor, and trace gas retrievals: recent updates evaluated with IASI case studies. *Atmos. Chem. Phys.* **13**, 6687–6711 (2013).
- Iacono, M. J. *et al.* Radiative forcing by long-lived greenhouse gases: calculations with the AER radiative transfer models. *J. Geophys. Res.* **113**, D13103 (2008).
- Stephens, G. L. *et al.* An update on Earth's energy balance in light of the latest global observations. *Nature Geosci.* **5**, 691–696 (2012).
- Manabe, S. & Wetherald, R. T. Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *J. Atmos. Sci.* **24**, 241–259 (1967).
- Stokes, G. M. & Schwartz, S. E. The Atmospheric Radiation Measurement (ARM) program: programmatic background and design of the cloud and radiation test bed. *Bull. Am. Meteorol. Soc.* **75**, 1201–1221 (1994).
- Clough, S. A. *et al.* Atmospheric radiative transfer modeling: a summary of the AER codes. *J. Quant. Spectrosc. Radiat.* **91**, 233–244 (2005).
- Turner, D. D. *et al.* Ground-based high spectral resolution observations of the entire terrestrial spectrum under extremely dry conditions. *Geophys. Res. Lett.* **39**, L10801 (2012).
- Peters, W. *et al.* An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proc. Natl Acad. Sci. USA* **104**, 18925–18930 (2007).
- Bruhwyler, L. M. *et al.* CarbonTracker-CH₄: an assimilation system for estimating emissions of atmospheric methane. *Atmos. Chem. Phys.* **14**, 8269–8293 (2014).
- Rienecker, M. M. *et al.* MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Clim.* **24**, 3624–3648 (2011).
- Gero, P. J. & Turner, D. D. Long-term trends in downwelling spectral infrared radiance over the U.S. Southern Great Plains. *J. Clim.* **24**, 4831–4843 (2011).
- Clothiaux, E. E. *et al.* Objective determination of cloud heights and radar reflectivities using a combination of active remote sensors at the ARM CART sites. *J. Appl. Meteorol.* **39**, 645–665 (2000).

25. Weatherhead, E. C. *et al.* Factors affecting the detection of trends: statistical considerations and applications to environmental data. *J. Geophys. Res.* **103** (D14), 17149–17161 (1998).
26. Haskins, R. D., Goody, R. M. & Chen, L. A statistical method for testing a general circulation model with spectrally resolved satellite data. *J. Geophys. Res.* **102**, 16563–16581 (1997).
27. Huang, Y. *et al.* Separation of longwave climate feedbacks from spectral observations. *J. Geophys. Res.* **115**, D07104 (2010).
28. Philipona, R., Dürr, B., Ohmura, A. & Ruckstuhl, C. Anthropogenic greenhouse forcing and strong water vapor feedback increase temperature in Europe. *Geophys. Res. Lett.* **32**, L19809 (2005).
29. Dessler, A. E. *et al.* Water-vapor climate feedback inferred from climate fluctuations, 2003–2008. *Geophys. Res. Lett.* **35**, L20704 (2008).
30. Haywood, J. M. *et al.* The roles of aerosol, water vapor and cloud in future global dimming/brightening. *J. Geophys. Res.* **116**, D20203 (2011).

Acknowledgements This material is based upon work supported by the US Department of Energy, Office of Science, Office of Biological and Environmental Research, Climate and Environmental Science Division, of the US Department of Energy under Award Number DE-AC02-05CH11231 as part of the Atmospheric System Research Program and the Atmospheric Radiation Measurement (ARM)

Climate Research Facility Southern Great Plains. We used resources of the National Energy Research Scientific Computing Center (NERSC) under that same award. I. Williams, W. Riley, and S. Biraud of the Lawrence Berkeley National Laboratory, and D. Turner of the National Severe Storms Laboratory also provided feedback. The Broadband Heating Rate Profile (BBHRP) runs were performed using Pacific Northwest National Laboratory (PNNL) Institutional Computing at PNNL, with help from K. Cady-Pereira of Atmospheric Environmental Research, Inc., L. Riihimaki of PNNL, and D. Troyan of Brookhaven National Laboratory.

Author Contributions D.R.F. implemented the study design, performed the analysis of all measurements from the ARM sites, and wrote the manuscript. W.D.C. proposed the study design and oversaw its implementation. P.J.G. is the AERI instrument mentor and ensured the proper use of spectral measurements and quality control. M.S.T. mentored the implementation of the study and oversaw its funding. E.J.M. and T.R.S. performed calculations and analysis to determine fair-weather bias. All authors discussed the results and commented on and edited the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to D.F. (drfeldman@lbl.gov).

METHODS

The methodology for this investigation focused on analysing time series of downwelling infrared spectra to determine the effects of CO₂ on these measurements and thereby to estimate its surface radiative forcing. The analysis also used temporally coincident measurements of atmospheric temperature and water vapour, retrievals of cloud occurrence, and assimilation estimates of CO₂ to construct simulated counterfactual measurements where CO₂ is held fixed.

Code and data availability. The measurement data sets used for this analysis are freely available through the ARM data repository (<http://www.arm.gov>). The radiative transfer codes are also freely available at <http://rtweb.aer.com>. CarbonTracker results were provided by NOAA/ESRL (<http://carbontracker.noaa.gov>). CarbonTracker-CH₄ results were provided by NOAA/ESRL (<http://www.esrl.noaa.gov/gmd/ccgg/carbontracker-ch4/>). The MERRA data used in this analysis are freely available for download at <ftp://goldsmr3.sci.gsfc.nasa.gov/data/s4pa/MERRA/MAI3CPASM.5.2.0/>. The Broadband Heating Rate Profile (BBHRP) data files, used to assess fair-weather bias, are freely available on the ARM BBHRP web page at <http://www.arm.gov/data/eval/24> under <http://dx.doi.org/10.5439/1163296> for the time-varying data stream and <http://dx.doi.org/10.5439/1163285> for the fixed CO₂ data stream. The computer routines used in this analysis will be made available upon request.

Schematic. The schematic (Extended Data Fig. 1) shows how CO₂ surface forcing is derived from differencing Atmospheric Emitted Radiance Interferometer (AERI) observations with a model calculation based on coincidental temperature and water vapour profiles. This produces a radiance difference (called a spectral residual) that is converted to flux units through a conversion factor calculated with a radiative-transfer model based on local meteorological conditions. The spectral residual, which is driven by observations, shows features only in the CO₂ absorption bands, meaning that this is a signal just from CO₂. The integration of forcing over frequency gives a broadband forcing term and contains a secular trend due largely to anthropogenic emissions.

Data sets. Spectra observed by the AERI instrument channel 1 were analysed over the period from 2000 through the end of 2010 at the ARM SGP and NSA sites. The AERI instrument contains two channels that cover a spectral range of 520–3,020 cm⁻¹ (19.2–3.3 μm) with a resolution of 0.5 cm⁻¹. Calibrated spectra are recorded every 8 min. Channel 1 observes the 520–1,800 cm⁻¹ spectral region, which covers at least 99.97% of the longwave CO₂ surface forcing. The absolute calibration is traceable to NIST standards³¹ and has been transferred from blackbodies with uncertainties better than 0.05 K (3σ). The resulting combined absolute uncertainty of each AERI spectrum is 1% (3σ)³². A summary of the AERI instrument's spectral stability can be seen in Extended Data Fig. 2, which shows the time series of effective laser wavenumber that is used to define the instrument's spectral scale. This time series exhibits a trend in the data of 4 ppm (relative) per decade around the nominal 15,799 cm⁻¹ laser wavenumber. Given that the uncertainty in the AERI spectral calibration is ±5 ppm (relative) (3σ), we have confidence that there is no statistically significant drift in the AERI spectral calibration over the course of the observation period (2000–2010).

Although the AERI record extends from 1995 to the present, temporally resolved estimates of the CO₂ concentration profiles from CT2011²⁰, which were necessary for validation of the radiative transfer model (see below), were only available beginning in 2000 and extending through 2010, resulting in an analysis period of 11 years. Overall, the AERI is a very robust instrument that inherently produces continuous, reliable operational data. Quality control for the spectra was achieved by ensuring that valid sky spectra were being observed by the instrument with the hatch open, and that unphysical radiance values, anomalously low variances in brightness temperature across the infrared window band (800–1,200 cm⁻¹) and scenes with substantial variability in the view of the hot blackbody were removed.

CO₂ profile best estimates were obtained from CT2011, which reported eight daily profiles in the vicinity of, and which were spatially interpolated to, the SGP and NSA sites at the European Centre for Medium-Range Weather Forecasting forecast model levels on a 40° latitude by 66° longitude grid covering North America and may have spin-up effects³³. The CT2011 assimilates data from many sources, including the average 12:00 to 16:00 local time CO₂ *in situ* measurement from the ARM Precision Carbon Dioxide Mixing Ratio System tower at 60 m altitude at SGP^{34,35}. Comparisons of CT2011 against aircraft profiles collected by the ARM Carbon Measurement Experiment³⁶ at SGP, which are not assimilated into CT2011, exhibit a bias of less than 0.5 ppm up to 3,600 m above sea level, with a standard deviation ranging from 1 ppm in winter to 3 ppm in summer. Between 2000 and 2011, the global CO₂ concentration at the surface increased from about 369 ppm to about 392 ppm, as measured by ARM-NOAA Earth Science Research Laboratory (ESRL) flasks, while the CO₂ concentrations at SGP and NSA increased at 1.95 ppm versus 1.98 ppm per decade, respectively. Time series of CT2011 profiles at SGP and fossil fuel and fire components are shown in Extended Data Fig. 3. Output from the CT2011 data set were used to create a baseline CO₂ value of

370 ppm from which surface-forcing calculations were determined. These outputs enabled us to demonstrate that spectral residuals derived from the Line-By-Line Radiative Transfer Model (LBLRTM) version 12.2 (ref. 18) with CT2011 had a lower rms difference than spectral residuals derived from the CT2011 baseline.

Temperature and humidity profiles were gathered from the ARM Balloon-Borne Sounding System (BBSS). At the SGP site, radiosondes were launched four times per day throughout the period of analysis, while at the NSA site, radiosondes were launched two times per day over the analysis period. Three different sounding technologies have been used at both locations. At the SGP site, the Vaisala RS-80 technology was used until April 2001; the Vaisala RS-90 was used between May 2001 and February 2005; and the Vaisala RS-92 was used from February 2005 to present. At the NSA site, RS-80 technology was used through April 2002, the RS90 was used from April 2002 until January 2005, and the RS-92 technology was used since January 2005. To account for the dry bias that has been exhibited particularly by the RS-80 sensors³⁷, we scaled the radiosonde-reported humidity profile by the ratio of the total column water retrieved from the microwave radiometer^{38,39} to the radiosonde-derived total column water. We estimate that the corrected radiosonde uncertainty is under 0.5 °C and 5% relative humidity (3σ), given the technology's stated precision and accuracy specifications. The change in radiosonde technology in the middle of the investigation could greatly complicate this analysis, and we would expect that a change in technology could introduce a jump-discontinuity in the forcing. The technology change occurred near the start of the analysis period in 2001, and the water-column amounts across the technology change differ at the *P* = 0.1 level for a two-sided *t*-test. However, we analysed the probability distribution function of the microwave radiometer to BBSS column water vapour both before and after the RS90 to RS92 change and did not find statistically significant differences (Extended Data Fig. 4).

Cloud clearing, which was essential to this analysis, was achieved through the ARM ARSCL Value-Added Product, which combines the Micropulse Lidar, Millimeter Wavelength Cloud Radar, and Vaisala and Belfort Ceilometers to produce six different cloud masks of varying sensitivity^{24,40}. We analysed spectra collected exclusively during clear-sky conditions, as identified by the absence of clouds in all of the coincident ARSCL masks.

Thermodynamic trends. Trends in the thermodynamic state at SGP and NSA are shown in Extended Data Fig. 5. The radiosonde-derived annually averaged least-squared trend in the clear-sky temperature profile at SGP showed an increase in spring in the lowest 2 km of the atmosphere, but decreases elsewhere. Lower-atmosphere water concentrations showed trends of opposite sign depending on season and altitude. At NSA, annually averaged clear-sky temperature profile trends were positive in the lower troposphere except in spring. Lower-tropospheric humidity also increased at NSA, except in the autumn. It is important to note that these thermodynamic trends are based on clear-sky conditions only and are distinct from all-sky trends.

Still, temperature trends can have a non-zero impact on spectral downwelling longwave radiation, which can be seen in Fig. 2e, where spectral trends at wavenumbers in the near-wings of the fundamental (ν₂) CO₂ band are negative. The cause of the negative trends arises from the nonlinear interactions between CO₂ and temperature trends, especially where there is a contrast in temperature trends between the boundary layer and the free troposphere under temperature inversion conditions, which are common at the NSA site. The effects of CO₂ and temperature generally produce separable impacts on the spectral downwelling longwave radiation, especially in the weak vibration-rotation bands that dominate the observed CO₂ surface radiative forcing. However, the effects interact nonlinearly near the centre of the fundamental (ν₂) CO₂ band²⁷. This nonlinearity complicates the separability of temperature and CO₂ trends over this narrow spectral range.

CarbonTracker-CH₄ and MERRA data use. All radiative transfer calculations performed in this analysis used CarbonTracker-CH₄ profile information. CarbonTracker-CH₄ (ref. 21) provides eight daily profiles on a 40° latitude by 60° longitude global grid of component contributions of CH₄ to the atmospheric concentration (a background component, a component due to coal and oil/gas production, a component due to animals, rice cultivation, and waste, a component due to wetlands, soil, oceans, and insects/wild animals, a component due to emissions from fires, and a component due to emissions from oceans). We sum those components and linearly interpolate them to the ARM SGP and NSA sites in space and time to coincide with the radiosonde observation.

We also used NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA)²² for O₃ profiles for all radiative transfer calculations. MERRA provides eight daily O₃ profiles at 42 pressure levels at 1.25° horizontal spacing and these were linearly interpolated to the ARM SGP and NSA sites in space and time to coincide with the radiosonde observation.

Radiative transfer. LBLRTM version 12.2 was used to simulate AERI spectra with and without CO₂ increases. LBLRTM was run on the 200 levels that are linearly spaced in log(pressure) onto which the radiosonde temperature and humidity

profiles have been interpolated. The conversion from spectral radiance to spectral flux was performed using spectrally dependent radiance-to-flux conversion factors determined through three-point quadrature flux calculations for each profile of atmospheric thermodynamic state corresponding to the AERI measurements following ref. 41. Histograms of the calculated distributions of the spectral conversion factor between radiance and flux are shown in Extended Data Fig. 6. These demonstrate that the conversion from radiance to flux depends on wavenumber, with stronger bands showing an Angular Distribution Model value very close to unity, indicating that the radiance is isotropic, and weaker bands having lower Angular Distribution Model values, indicating that the radiance is anisotropic.

To determine the contribution of CO₂ to surface forcing, we removed the signal associated with the other radiatively active constituents. To do this, we performed a radiative-transfer calculation using LBLRTM with observed temperature and water-vapour profiles from radiosondes (with water vapour scaled by microwave radiometer column water retrievals) and with a fixed CO₂ concentration set to 370 ppm (the CO₂ atmospheric concentration at the beginning of the analysis period) for cloud-free conditions when the AERI instrument produced calibrated spectra. We differenced these calculations against the actual AERI spectral measurements to produce a time-series of residual spectra between 2000 and the end of 2010. Since the fixed CO₂ concentration is the primary difference between the inputs to each calculation and the actual time-varying atmospheric states, the residual spectra contains the signal of changing CO₂. We integrated the spectra over frequency (wavenumber) and converted radiance to flux, based on the quadrature calculations described above, to produce a forcing value at each time step.

We used the Rapid Radiative Transfer Model¹⁸ to perform the calculations of fair-weather bias, as described below. It calculates lower radiative forcing values than line-by-line models for the same atmospheric state¹⁴.

Error budget. The error budget for each surface-forcing value contains contributions from several sources, including radiometric noise, instrument spectral and radiometric calibration, and residual removal. Instrument radiometric calibration is achieved through a suite of diagnostics, including detector nonlinearity characterization, electronics calibration, field-of-view testing, and routine views of the instrument blackbodies. The achieved radiometric uncertainty is better than 1% (3 σ) of ambient radiance for a single observation and decreases with the number of observations. The instantaneous noise-effective spectral radiance specification for the AERI instrument is $<0.2 \text{ mW m}^{-2} \text{ sr}^{-1} \text{ per cm}^{-1}$ for 670–1,400 cm^{-1} , for the rms of a 2-min ambient blackbody view. Spectral calibration is achieved through routine comparisons with stable atmospheric lines and has an uncertainty of better than 0.08 cm^{-1} (corresponding to 5 ppm of the laser wavenumber) (3 σ). Residual removal uncertainty consists of two subcomponents: (1) uncertainty in the atmospheric structure as provided by the BBSS and (2) uncertainty in the spectroscopy that informs LBLRTM. The former is derived from the radiosonde uncertainty propagated in spectral space with LBLRTM; the latter is derived from the HITRAN database uncertainty codes¹⁰. The effects of spectroscopic uncertainty are derived from forcing estimates with LBLRTM calculations where line parameters are modified according to stated HITRAN error codes, which report uncertainties in intensity, half-width, temperature dependence, and pressure shift for each line.

Trend determination. Both spectral trends, as shown in Fig. 2 and broadband trends, as shown in Fig. 4, are determined from a least-squares regression using Matlab's "polyfit" (<http://www.mathworks.com/help/matlab/ref/polyfit.html>) function of the estimate values based on clear-sky conditions only. Uncertainty in the trend was calculated from the formulae in Weatherhead *et al.*²⁵. The trend in seasonal amplitude was determined from a least-squares regression from the peaks of the monthly averaged time-series of the forcing.

Fair-weather bias. One key question associated with the methods concerns the potential for fair-weather bias in this approach due to the screening of data based on the ARSCL cloud masks. A bias could occur if there is a relationship between surface forcing and cloud cover that is distinct from, and would alter the trends in, what was found under clear-sky conditions. To address this question, we have undertaken a set of calculations that follow the BBHRP⁴² calculations based on the Radiatively Important Parameters Best Estimate (RIPBE)⁴³ Value Added Products (see <http://www.arm.gov/data/eval/24> for details). This product contains broadband profiles of fluxes and heating rates throughout the atmosphere. We have recalculated BBHRP profiles at SGP for 2010 based on fixed CO₂ concentrations and compared those to the original BBHRP profiles (that contained time-varying CO₂) and subset the data that were identified as clear-sky in RIPBE (through the ARSCL flags) in order to estimate this bias. The results are shown in Extended Data Fig. 7 and indicate that the all-sky surface forcing is 0.05 W m^{-2} less than the clear-sky forcing, which represents a 25% difference in the two quantities. This finding is expected because there is non-negligible overlap between cloud absorption, which is broadband, and CO₂ absorption features, so clouds mask the forcing from CO₂. We note that the difference in clear-sky and all-sky surface radiative forcing values is not statistically significant and that the extent of the fair-weather bias depends on

the occurrence frequency of cloudy conditions, which varies year-to-year at the SGP and NSA sites.

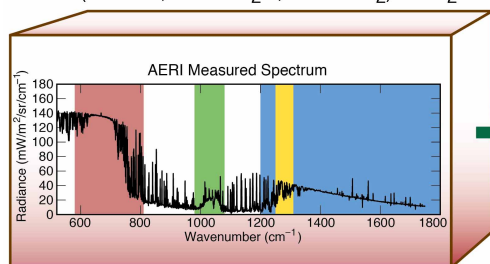
Surface versus TOA forcing. In general, several studies have found that there is a complicated relationship between surface and top-of-atmosphere (TOA) forcing, depending on the atmospheric thermodynamic and latent heating structure, with the former being less than the latter^{15,44}. This is further underscored by other findings for an increase in CO₂ from 287 to 369 parts per million by volume (ppmv) which shows a forcing at the surface of 0.57 W m^{-2} and a forcing at 200 hPa of 1.92 W m^{-2} from line-by-line models⁴⁵.

To specifically address the relationship between forcing at the surface and TOA at the ARM sites, we have also undertaken a set of calculations that use the BBHRP infrastructure described above to demonstrate the relationship between surface and TOA flux changes. Results are shown in Extended Data Fig. 8, and indicate that magnitudes of the flux changes at the surface and TOA are correlated, although the TOA flux changes are larger because the downwelling surface radiation of the former quantity is distributed into sensible and latent heat. We note that the flux change at the surface and TOA is calculated by differencing fluxes with time-varying CO₂ minus fluxes with fixed CO₂ at each altitude, and the decreased emission to space and increased emission to the surface leads to flux changes of opposite sign for TOA and surface flux changes, respectively.

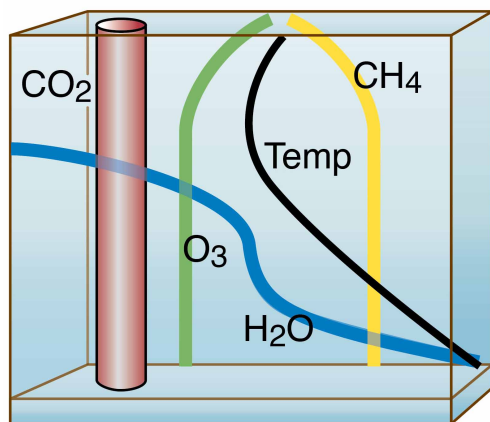
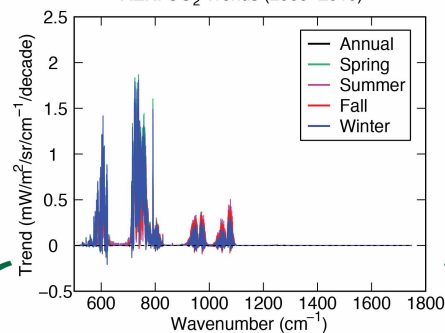
Vertical sensitivity. Using LBLRTM, we calculated the vertical sensitivity of broadband longwave surface forcing due to perturbations in the atmospheric state profile. These calculations were based on model atmospheres, including the US standard, the tropical, the mid-latitude summer, the mid-latitude winter, the sub-Arctic summer, and the sub-Arctic winter atmospheres, which span a broad range of atmospheric thermodynamic states⁴⁶. We separately perturbed the temperature by 1°K, 1% water vapour, 10 ppm CO₂, 10% O₃, and 1 ppm CH₄ profiles for a 1-km-thick layer of the atmosphere from the surface to 20 km. Results are shown in Extended Data Fig. 9 and indicate that, with the exception of O₃, contributions to surface forcing are dominated by the bottom 5 km of the atmosphere, regardless of the thermodynamic state. O₃ perturbations have a more complicated vertical structure because O₃ is partially transparent in the troposphere and its concentration peaks in the stratosphere, meaning that upper-level O₃ concentration perturbations at the 10% level can have a modest impact on the surface energy balance.

31. Best, F. A. *et al.* Traceability of absolute radiometric calibration for the Atmospheric Emitted Radiance Interferometer (AERI). In *Conf. on Characterization and Radiometric Calibration for Remote Sensing (Space Dynamics Laboratory, Utah State Univ., 15–18 Sept)* <http://www.calcon.sdl.usu.edu/conference/proceedings> (2003).
32. Knuteson, R. O. *et al.* Atmospheric Emitted Radiance Interferometer. Part II: Instrument performance. *J. Atmos. Ocean. Technol.* **21**, 1777–1789 (2004).
33. Masarie, K. A. *et al.* Impact of CO₂ measurement bias on CarbonTracker surface flux estimates. *J. Geophys. Res.* **116**, D17305 (2011).
34. Bakwin, P. S., Tans, P. P., Zhao, C., Ussler, W. & Quesnell, E. Measurements of carbon dioxide on a very tall tower. *Tellus B.* **47**, 535–549 (1995).
35. Bakwin, P. S., Tans, P. P., Hurst, D. F. & Zhao, C. Measurements of carbon dioxide on very tall towers: results of the NOAA/CMDL program. *Tellus B.* **50**, 401–415 (1998).
36. Biraud, S. C. *et al.* A multi-year record of airborne CO₂ observations in the US Southern Great Plains. *Atmos. Technol.* **6**, 751–763 (2013).
37. Wang, J. *et al.* Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—application to TOGA COARE data. *J. Atmos. Ocean. Technol.* **19**, 981–1002 (2002).
38. Liljegren, J. C. in *Microwave Radiometry and Remote Sensing of the Earth's Surface and Atmosphere* (eds Pampaloni, P. & Paloscia, S.) 433–443 (VSP Press, 1999).
39. Cimini, D., Westwater, E. R., Han, Y. & Keilm, S. J. Accuracy of ground-based microwave radiometer and balloon-borne measurements during the WVIOP2000 field experiment. *IEEE Trans. Geosci. Rem. Sens.* **41**, 2605–2615 (2003).
40. Clothiaux, E. E. *et al.* The ARM Millimeter Wave Cloud Radars (MMCRs) and the Active Remote Sensing of Clouds (ARSCL) Value Added Product (VAP) DOE Tech. Memo. ARM VAP-002.1, https://www.arm.gov/publications/tech_reports/arm-vap-002-1.pdf (US Department of Energy, 2001).
41. Li, J. Gaussian quadrature and its application to infrared radiation. *J. Atmos. Sci.* **57**, 753–765 (2000).
42. Mlawer, E. J. *et al.* The broadband heating rate profile (BBHRP) VAP. *Proc. 12th ARM Sci. Team Meet. ARM-CONF-2002* http://www.arm.gov/publications/proceedings/conf12/extended_abs/mlawer-ej.pdf (US Department of Energy, 2002).
43. McFarlane, S., Shippert, T. & Mather, J. Radiatively Important Parameters Best Estimate (RIPBE): an ARM value-added product. DOE Tech. Rep. SC-ARM/TR-097 https://www.arm.gov/publications/tech_reports/doe-sc-arm-tr-097.pdf (US Department of Energy, 2011).
44. Andrews, T., Forster, P. M., Boucher, O., Bellouin, N. & Jones, A. Precipitation, radiative forcing and global temperature change. *Geophys. Res. Lett.* **37**, L14701 (2010).
45. Collins, W. D. *et al.* Radiative forcing by well-mixed greenhouse gases: estimates from climate models in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). *J. Geophys. Res.* **111**, D14317 (2006).
46. Anderson, G. P. *et al.* AFGL atmospheric constituent profiles (0–120 km). AFGL-TR 86-0110, <http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA175173> (Hanscom Air Force Base, Air Force Geophysics Laboratory, 1986).

AERI-F(RAOB T, RAOB H₂O, Fixed CO₂) = CO₂ Forcing

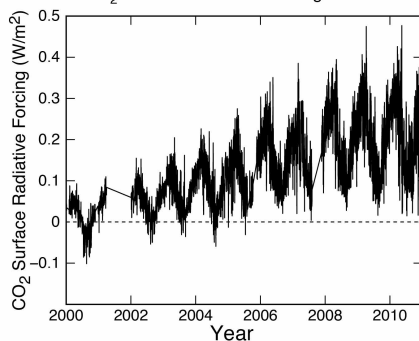


AERI CO₂ Trends (2000–2010)

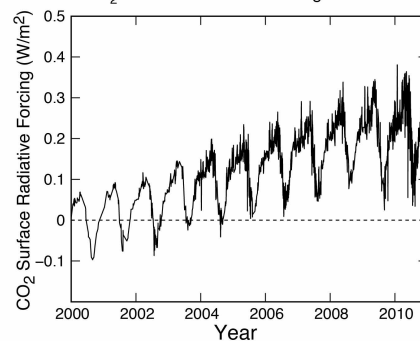


Atmospheric State

SGP CO₂ Surface Radiative Forcing Time-Series

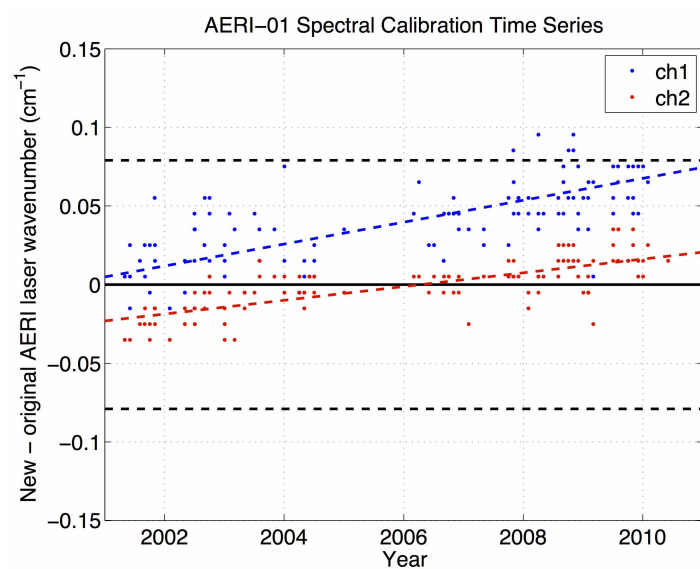


NSA CO₂ Surface Radiative Forcing Time-Series

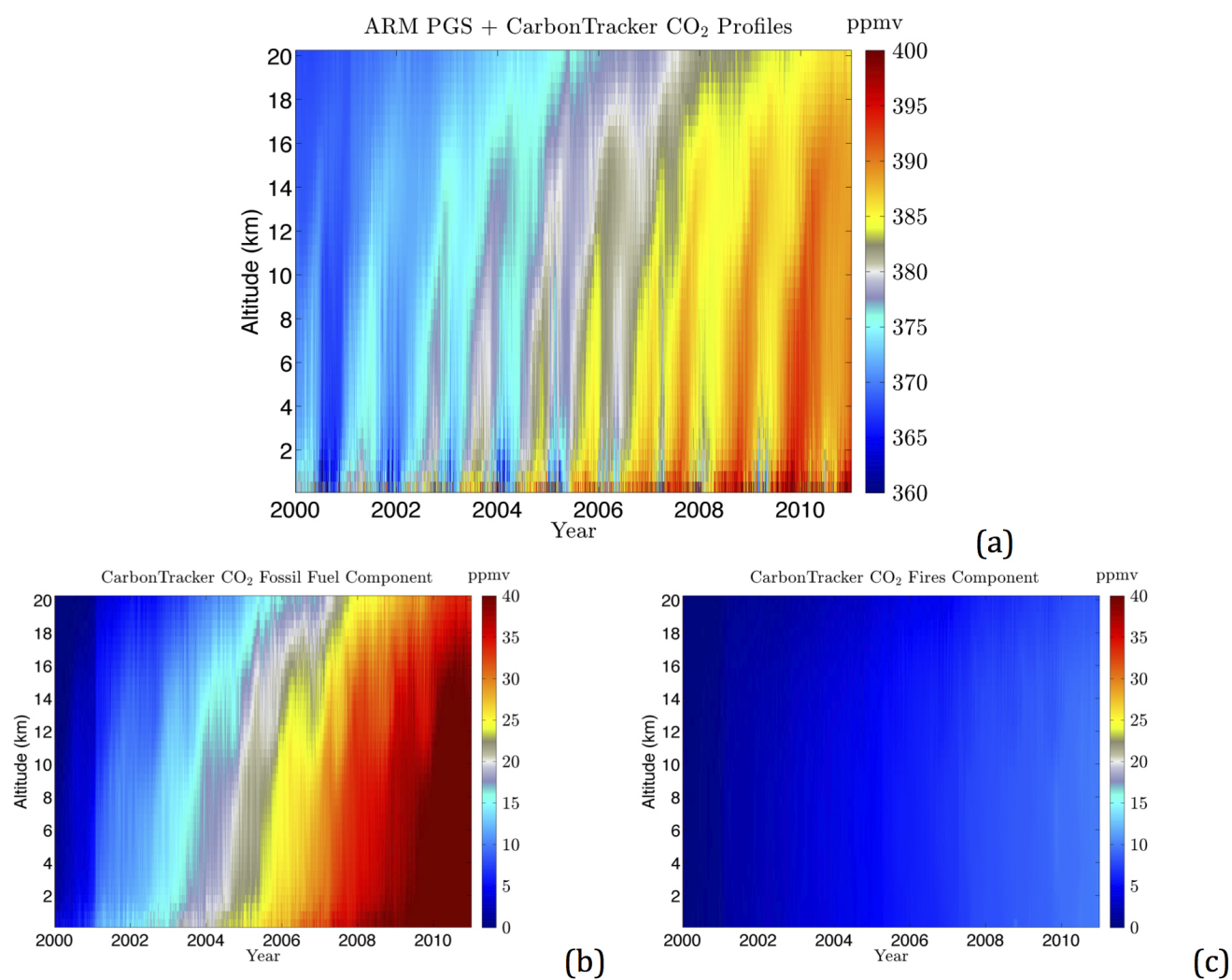


Time-series shows steady increase in CO₂ surface forcing and many higher-frequency features

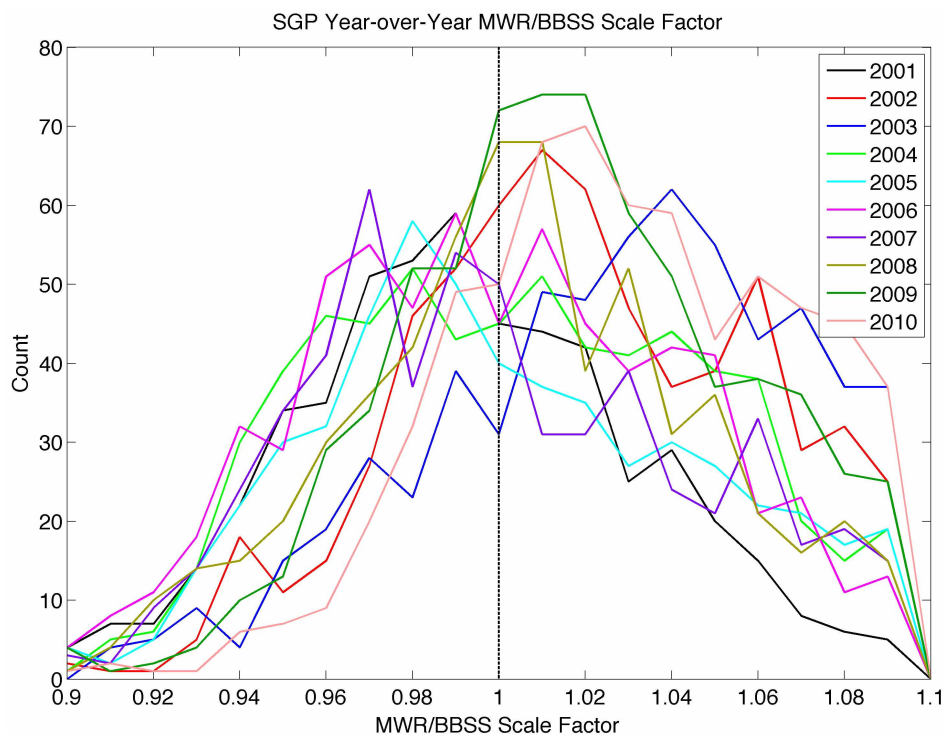
Extended Data Figure 1 | Schematic. Schematic of the derivation of surface forcing from AERI observations and calculations based on the atmospheric structure.



Extended Data Figure 2 | AERI instrument stability. Time series of the AERI-instrument-derived laser wavenumber around a nominal frequency of $15,799 \text{ cm}^{-1}$.



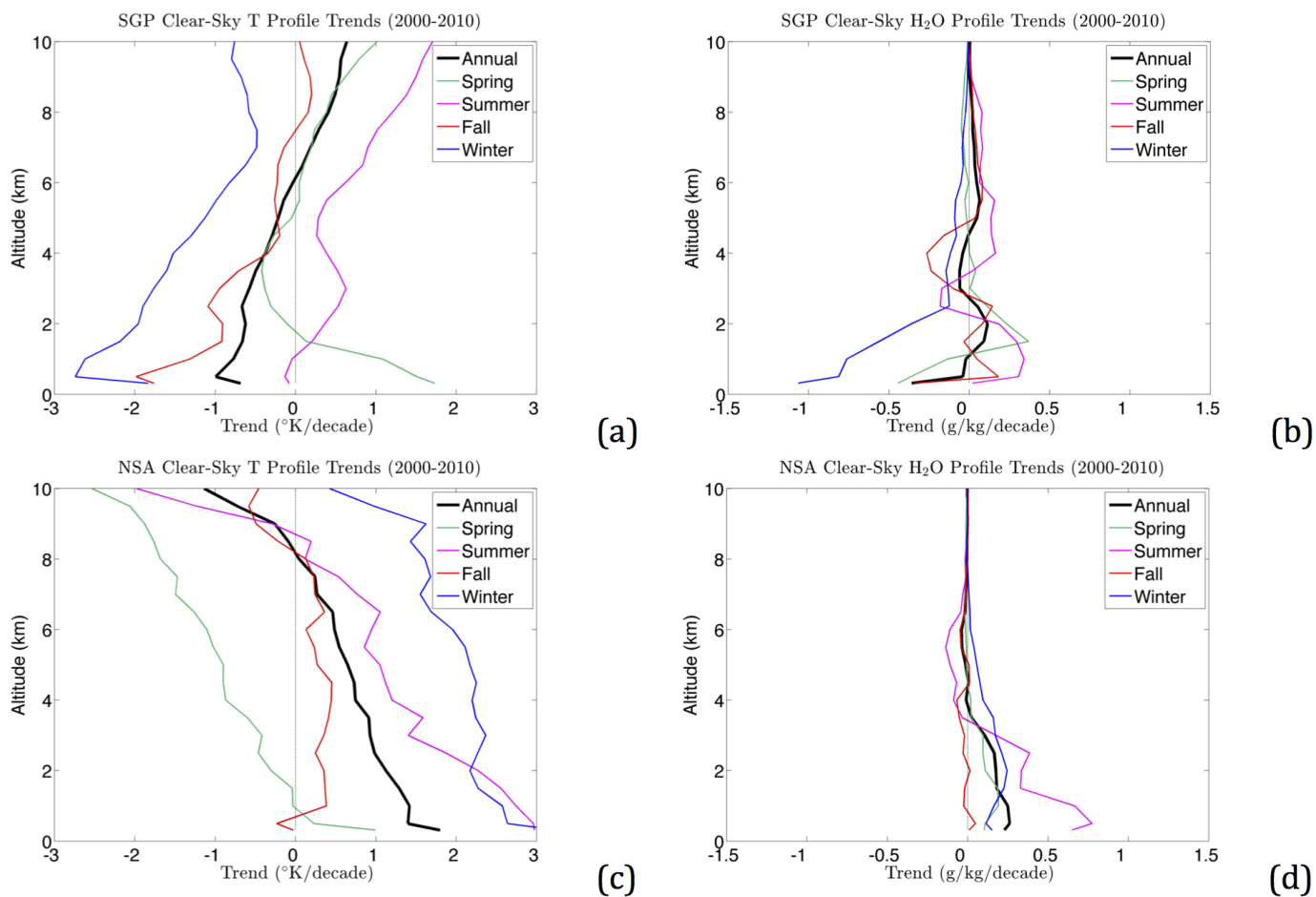
Extended Data Figure 3 | CarbonTracker profiles. **a**, CT2011 profile time series of CO₂ at the SGP site. **b**, CT2011 fossil fuel component of the CO₂ profile. **c**, CT2011 biomass burning component of the CO₂ profile. PGS, the ARM Precision Gas System Carbon Dioxide Mixing Ratio System.



Extended Data Figure 4 | Microwave radiometer radiosonde scaling.

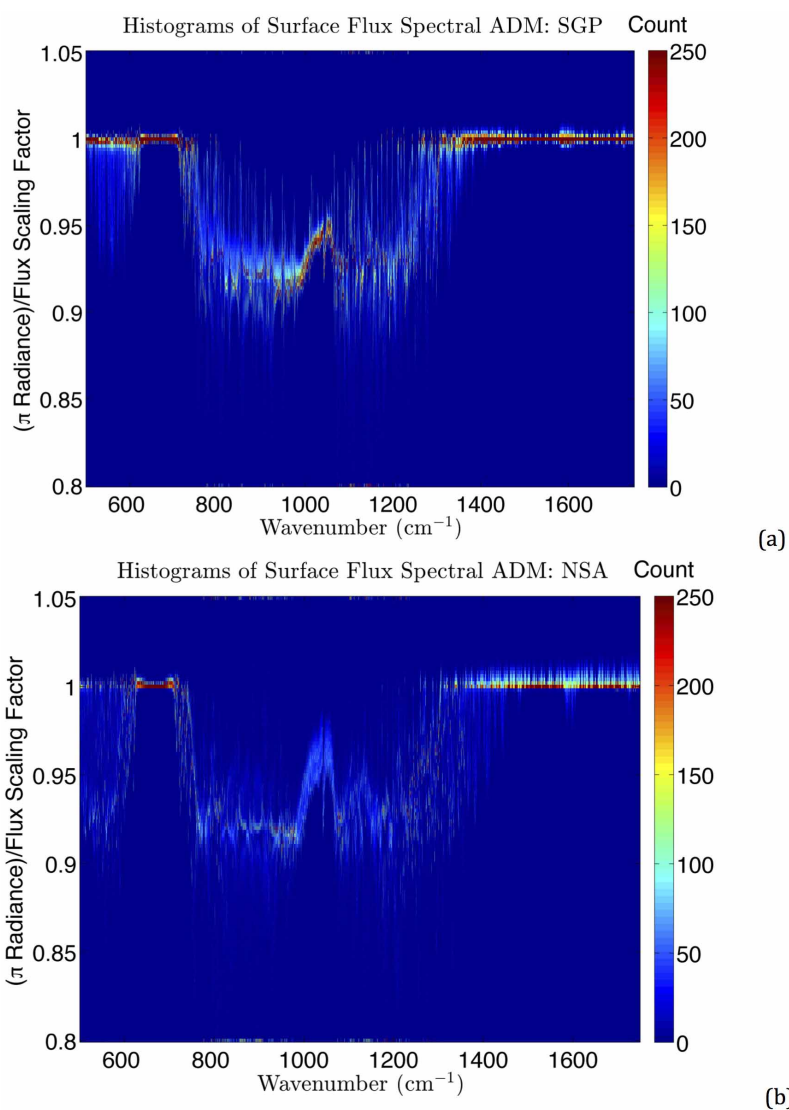
Distribution of microwave radiometer (MWR) precipitable water vapour to the precipitable water vapour derived from radiosondes for each year of the

investigation at the ARM SGP site. Each count corresponds to the scaling between a collocated radiosonde and microwave radiometer retrieval.



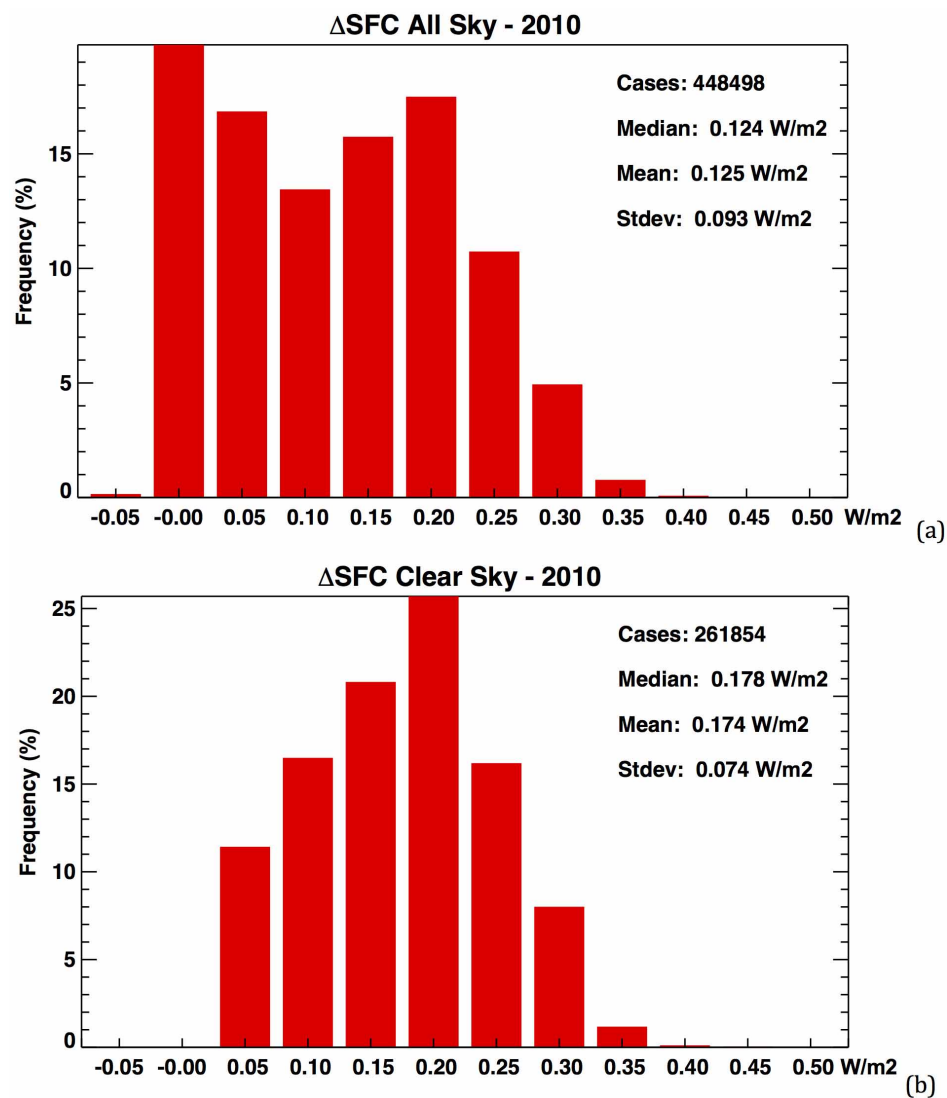
Extended Data Figure 5 | Thermodynamic trends. **a**, Annual and seasonal clear-sky temperature (T) profile trends derived from radiosondes and ARSCL data for cloud-clearing at SGP from 2000 to 2010. **b**, Same as **a** but for

water vapour (H_2O) profile trends. **c**, As for **a** but temperature profile trends at NSA. **d**, As for **b** but for water vapour profile trends (in grams of water vapour per kilogram of air per decade) at NSA.

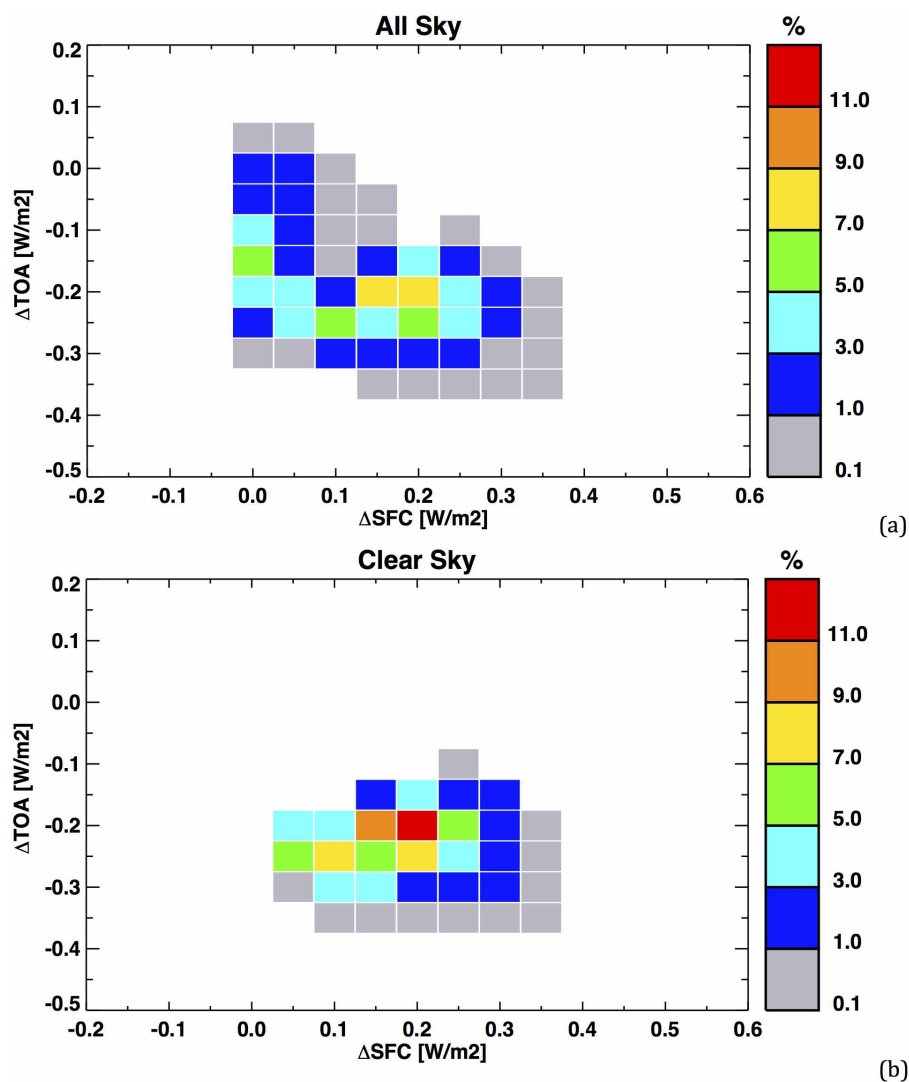


Extended Data Figure 6 | Conversion from radiance to flux. Histogram of zenith radiance to flux spectral conversion for AERI channel 1 spectral channels based on LBLRTM calculations based on the thermodynamic profiles

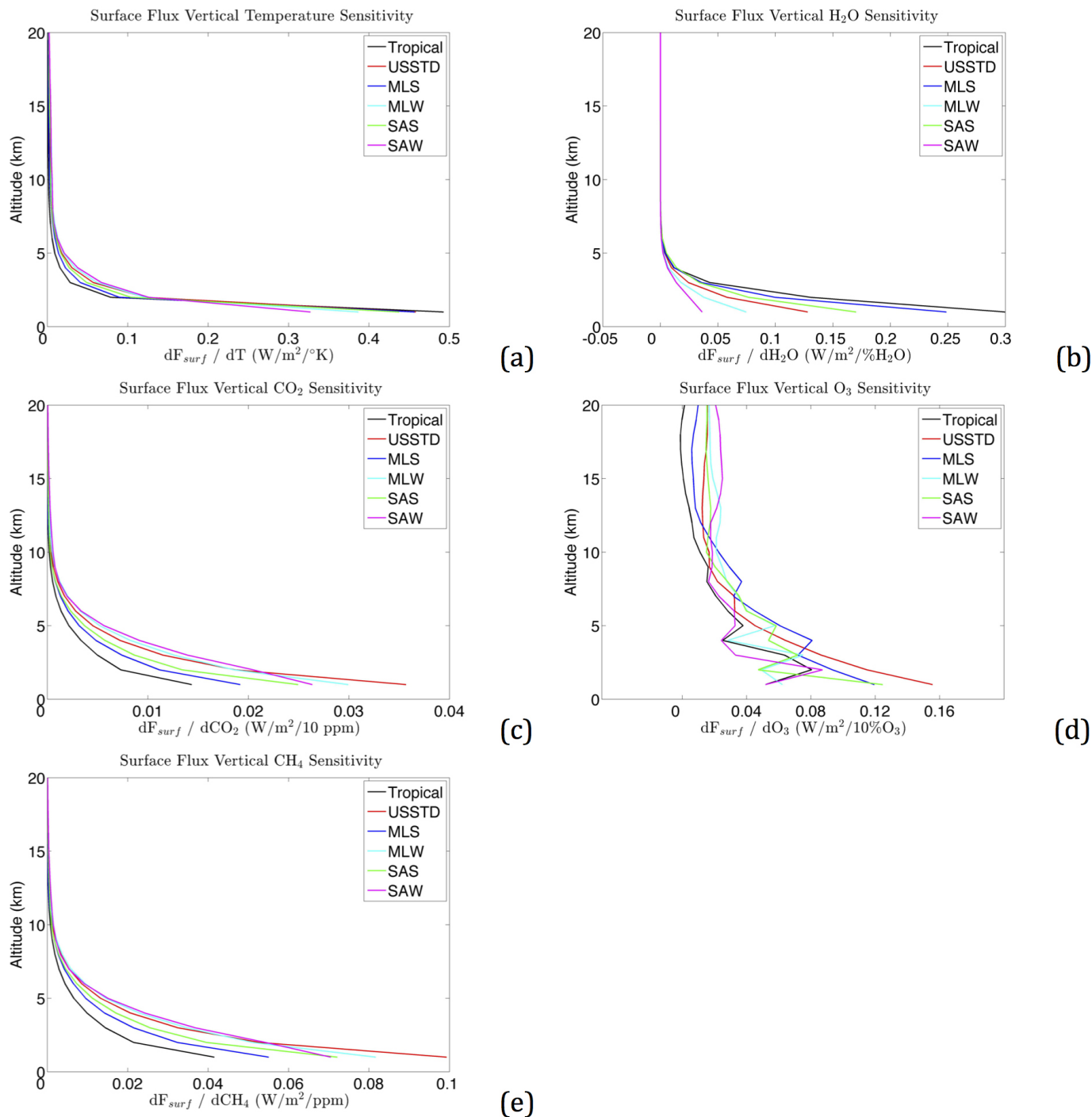
from the ARM SGP site from 2000 to 2010. **b**, As for **a** but for the NSA site. ADM, Angular Distribution Model.



Extended Data Figure 7 | Fair-weather bias. **a**, Histogram of the difference in flux between BBHRP calculations with time-varying CO₂ and calculations where CO₂ = 370 ppmv for all profiles at 30-min resolution during 2010 at SGP. **b**, As for **a** but for the subset of data identified by the ARSCL as clear-sky.



Extended Data Figure 8 | TOA and surface fluxes. **a**, Occurrence frequency (in per cent) plot of tropopause versus surface forcing based on BBHRP calculations with time-varying CO₂ and where CO₂ = 370 ppmv for all profiles at 30-min resolution during 2010 at SGP for all-sky conditions as identified by ARSCL flags. **b**, As for **a** but for clear-sky conditions.



Extended Data Figure 9 | Surface flux sensitivity to atmospheric profiles.

a. The sensitivity of the surface radiative flux (F_{surf}) to the level of a $1^\circ K$ perturbation in temperature for different model atmospheres including tropical, US standard (USSTD), mid-latitude summer (MLS), mid-latitude

winter (MLW), sub-Arctic summer (SAS), and sub-Arctic winter (SAW)⁴⁶.

b. As for **a** but level perturbations are given as percentage H_2O . **c.** As for **a** but level perturbations are 10 ppm CO_2 . **d.** As for **a** but level perturbations are 10% O_3 . **e.** As for **a** but level perturbations are 1 ppm CH_4 .

Long-term decline of the Amazon carbon sink

A list of authors and their affiliations appears at the end of the paper

Atmospheric carbon dioxide records indicate that the land surface has acted as a strong global carbon sink over recent decades^{1,2}, with a substantial fraction of this sink probably located in the tropics³, particularly in the Amazon⁴. Nevertheless, it is unclear how the terrestrial carbon sink will evolve as climate and atmospheric composition continue to change. Here we analyse the historical evolution of the biomass dynamics of the Amazon rainforest over three decades using a distributed network of 321 plots. While this analysis confirms that Amazon forests have acted as a long-term net biomass sink, we find a long-term decreasing trend of carbon accumulation. Rates of net increase in above-ground biomass declined by one-third during the past decade compared to the 1990s. This is a consequence of growth rate increases levelling off recently, while biomass mortality persistently increased throughout, leading to a shortening of carbon residence times. Potential drivers for the mortality increase include greater climate variability, and feedbacks of faster growth on mortality, resulting in shortened tree longevity⁵. The observed decline of the Amazon sink diverges markedly from the recent increase in terrestrial carbon uptake at the global scale^{1,2}, and is contrary to expectations based on models⁶.

The response of the Earth's land surface to increasing levels of atmospheric CO₂ and a changing climate provide important feedbacks on future greenhouse warming^{6,7}. One of the largest ecosystem carbon pools on Earth is the Amazon forest, storing around 150–200 Pg C in living biomass and soils⁸. Earlier studies based on forest inventories in the Amazon Basin showed the tropical forest here to be acting as a strong carbon sink with an estimated annual uptake of 0.42–0.65 Pg C yr⁻¹ for 1990–2007, around 25% of the residual terrestrial carbon sink^{3,4}. There is, however, substantial uncertainty as to how the Amazon forest will respond to future climatic and atmospheric composition changes. Some earlier modelling studies predicted a large-scale dieback of the Amazon rainforest⁹, while more recent studies predict a carbon sink well into the twenty-first century due to a CO₂ fertilization effect⁶. The realism of such model predictions remains low owing to uncertainty associated with future climate and vegetation responses^{6,7} in particular changes in forest dynamics^{5,10,11}. Thus, direct observations of tropical tree responses are crucial to examine what changes are actually occurring and what to expect in the future. Here we analyse the longest and largest spatially distributed time series of forest dynamics for tropical South America.

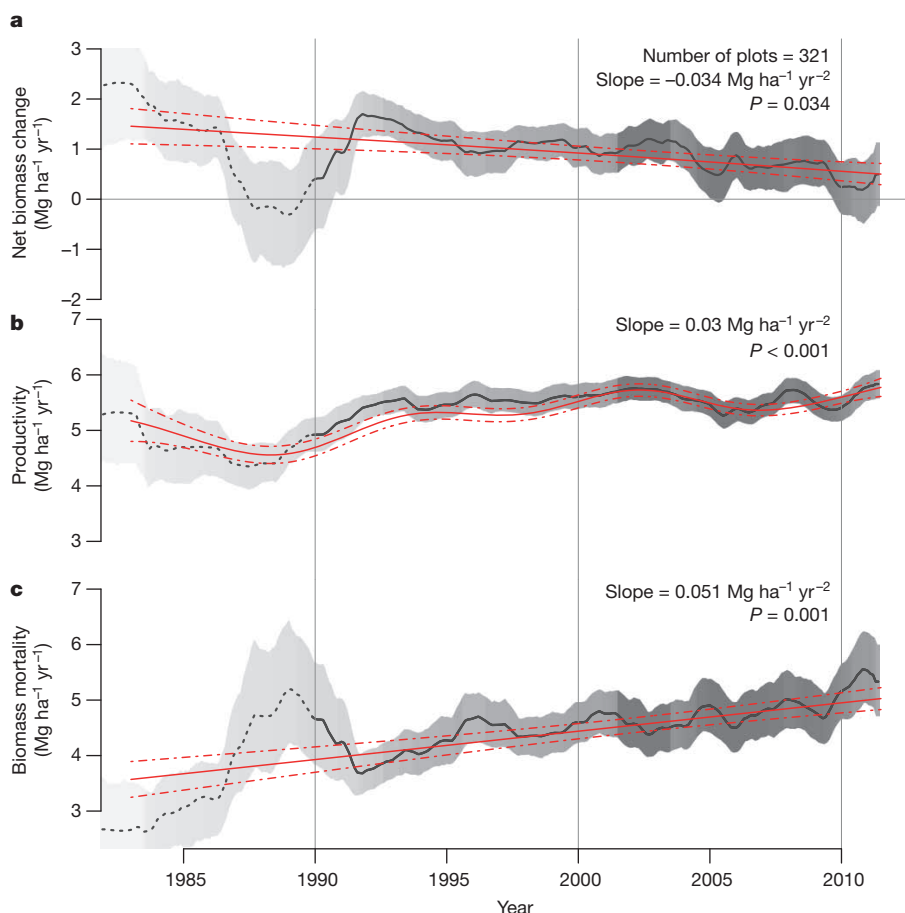


Figure 1 | Trends in net above-ground biomass change, productivity and mortality across all sites. a–c, Black lines show the overall mean change up to 2011 for 321 plots (or 274 units) weighted by plot size, and its bootstrapped confidence interval (shaded area). The red lines indicate the best model fit for the long-term trends since 1983 using general additive mixed models (GAMM), accounting explicitly for differences in dynamics between plots (red lines denote overall mean, broken lines denote s.e.m.). Alternative analyses of subsets of plots that were all continuously monitored throughout shorter time intervals confirm that the observed trends are not driven by temporal changes in individual sample plot contributions (Extended Data Fig. 3). Estimated long-term (linear) mean slopes and significance levels are indicated, and are robust with regard to the statistical approach applied (that is, parametric or non-parametric, see Methods). Shading corresponds to the number of plots that are included in the calculation of the mean, varying from 25 plots in 1983 (light grey) to a maximum of 204 plots in 2003 (dark grey). The uncertainty and variation is greater in the early part of the record owing to relatively low sample size (see Extended Data Fig. 4).

Our analysis is based on 321 inventory plots lacking signs of recent anthropogenic impacts from the RAINFOR network⁴ and published plots. The sites are distributed throughout the Amazon basin and cover all major forest types, soils and climates (Extended Data Fig. 1). For each plot (mean size 1.2 ha) all trees with stem diameter greater than 100 mm were identified, and allometric equations applied to convert tree diameter, height and wood density to woody biomass or carbon⁸. Net biomass change was estimated for each census interval as the difference between standing biomass at the end and the beginning of the interval divided by the census length. We also derived forest woody productivity (hereafter termed productivity) from the sum of biomass growth of surviving trees and trees that recruited (that is, reached a diameter ≥ 100 mm), and mortality from the biomass of trees that died between censuses, allowing for census-interval effects (see Methods). Plots were measured on average five times and the mean measurement period was 3 years. For analysis purposes small plots were aggregated to leave 274 distinct units. We report trends since 1983, the first year with measurements for 25 plots, up to mid-2011.

Our data show that mature forests continued to act as a biomass sink from 1983 to 2011.5, but also reveal a long-term decline in the net rate of biomass increase throughout the census period (Fig. 1a). The decline

in net biomass change is due to a strong long-term increase in mortality rates (Fig. 1c), and occurred despite a long-term increase in productivity (Fig. 1b). While mortality increased throughout the period, productivity increases have recently stalled showing no significant trend since 2000 (Extended Data Fig. 3). These time trends are based on a varying set of plots over time (Extended Data Fig. 4), but this site-switching does not alter the results (see Methods). The observed trends also emerge from a separate plot-by-plot analysis (Fig. 2), with increases in mortality exceeding productivity gains by approximately two to one. Trends are rarely significant at the individual plot level owing to the stochastic nature of local forest dynamics, but the mean slopes of net change, productivity and mortality all differ significantly from zero. Changes in forest dynamics were not geographically limited to a particular area, but occurred throughout the lowland South American tropics (Fig. 2). While rates of change vary depending on the precise plot set, time window and analytical approach used, the trends remain robust (Figs 1, 2 and Extended Data Fig. 3).

Artefactual explanations have previously been offered to explain trends in biomass dynamics from plot measurements^{12,13}. Principally, it has been suggested that reported net biomass increases⁴ could be driven by recovery of forests from local disturbances¹². However, contrary to

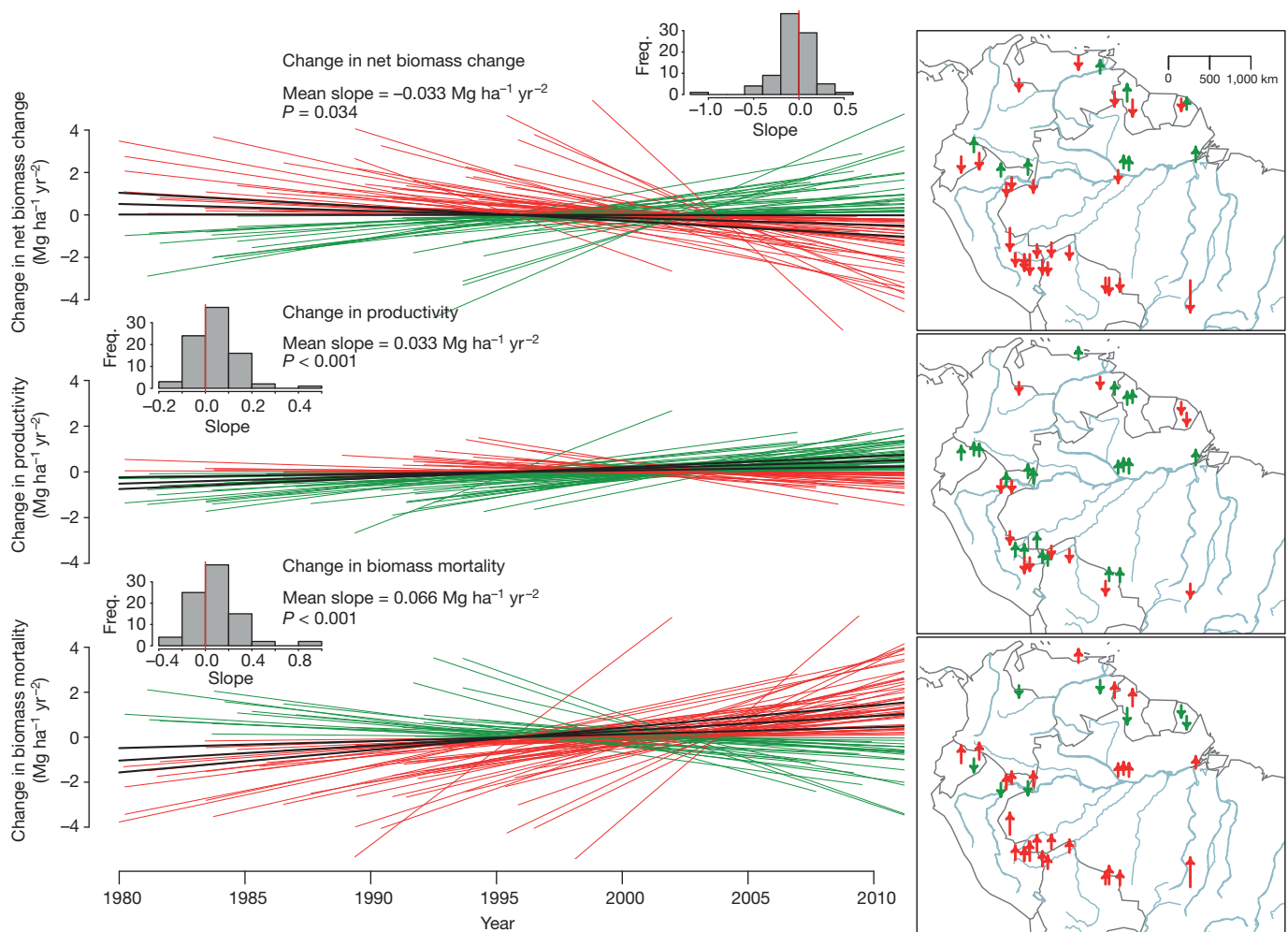


Figure 2 | Annual change in net above-ground biomass change, productivity and mortality for individual sites. The lines in the left-hand panels show the long-term rate of change for 117 plots (or 87 units), estimated using linear regressions weighted by census-interval length and for display purposes centred around zero. This analysis includes only plots that were monitored for at least 10 years and contained three or more census intervals with at least one in the 1990s and one in 2000s. Red lines indicate long-term trends that negatively affect biomass stocks (for example, decreasing net

change, increasing losses) and green lines indicate trends that positively affect biomass stocks (for example, increasing productivity). Bold black lines indicate the mean slope across all plots and confidence intervals (2.5–97.5 percentiles). Insets in the left panels show the frequency distribution of the slopes, with the mean slope and P value for t -test of difference from no slope. The maps show the location of the sites, and the colour and arrow length indicate the sign and magnitude of the slope, with adjacent plots joined into a single site for display purposes.

observations from recovering neotropical forests¹⁴ and successional studies¹⁵, the plots have collectively experienced increased biomass growth (Fig. 1), accelerated stem recruitment and death (Extended Data Fig. 6), and net biomass change is positively related to changes in stem numbers, but not in wood density (Fig. 3b, c). It is thus unlikely that the overall patterns would be driven by recovery from disturbances. Alternatively, increases in mortality have been proposed to arise due to biased selection of plots in mature forest patches, which over time accumulate disturbances and so decline in biomass¹³. The fact that forests and trees have continued to get bigger (Extended Data Fig. 5a) is contrary to this explanation. In addition, if this were driving the network-wide pattern, then the observed trends should disappear if data are reanalysed using only the first interval of each plot, but instead they persist. In summary, the data suggest that trends are unlikely to be caused by artefactual explanations of forests recovering from disturbances or selection of mature forest patches (see Supplementary Information for a more complete exploration of these potential biases).

The factors driving the observed long-term changes remain unclear. The levelling off of productivity in the most recent decade (Fig. 1b and Extended Data Fig. 3f) could be due either to a relaxation of the growth stimulus itself, or to the onset of a counteracting factor depressing growth rates. The recent demonstration of Amazon-wide carbon sink suppression during a drought year¹⁶ indicates one possible driver. Tropical drought is also often associated with higher temperatures, which may further contribute to reducing productivity¹⁷ and carbon uptake¹⁸. The past decade in Amazonia has seen several droughts¹⁹ and warming²⁰, which coincide closely with the stalling productivity across Amazon forests.

The increased rate of biomass mortality is driven by an increasing number of trees dying per year (Extended Data Fig. 6c) rather than an increase in the size of the dying trees (Extended Data Fig. 5c). Several mechanisms may explain this increase in loss of biomass due to tree mortality, with recent climate events being an obvious candidate. The plot data clearly show short-term peaks in the size of dying trees during the anomalously dry years 2005 and 2010 (Extended Data Fig. 5c). These are consistent with results from rainfall exclusion experiments in Amazonia^{21,22} and observations⁴ showing that large tropical trees are vulnerable to drought stress. However, our data lack the signature expected if drought were the dominant long-term driver of the increasing loss of biomass due to mortality in Amazonia. That is, there has been no long-term change in the size of dead trees (Extended Data Fig. 5c), living trees have continued to get bigger (Extended Data Fig. 5a), and the increase in stem mortality predates the drought of 2005 (Extended Data Fig. 6c).

Alternatively, the increased productivity may have accelerated tree life cycles so that they now die younger. Large stature is associated with size-related hydraulic²³ and mechanical failure²⁴, reproductive costs²⁵ and photosynthetic decline²³. Faster growth exposes trees to these size-related risks earlier, as evidenced by tree ring data suggesting that faster growth shortens lifespans^{26,27}, and by experimental data showing early onset of reproduction under increased CO₂ (ref. 28). The observed long-term acceleration in stem mortality rates and the plot-level association between productivity and the strength of the increase in biomass loss due to mortality (Extended Data Fig. 8b) are consistent with such a mechanism. While demographic feedbacks are not explicitly included in dynamic global vegetation models¹⁰, our results suggest that they could in fact influence the capacity of forests to gain biomass²⁹, with transient rates of ecosystem net carbon accumulation highly sensitive to even small changes in carbon turnover times¹⁰.

Finally, we put our results in a global perspective. According to global records, the land carbon sink has increased since the mid-1990s (refs 1, 2). While tropical land contributed significantly to this global sink during the 1980s and 1990s, our results show that the total net carbon sink into intact Amazon live biomass then decreased by 30% from 0.54 Pg C yr⁻¹ (confidence interval 0.45–0.63) in the 1990s to 0.38 Pg C yr⁻¹ (0.28–0.49) in the 2000s (see Methods). If our findings for the Amazon are representative for other tropical forests, and if below-ground pools have responded in the same way as above-ground biomass (AGB), then an apparent divergence emerges between a strengthening global terrestrial sink on one hand^{1,2} and a weakening tropical sink on the other. However, from an atmospheric perspective we also note that some of the effects of the Amazon changes are yet to be observed, as little of the carbon resulting from increased mortality is immediately released into the atmosphere³⁰. Instead, dead trees decay slowly, with a fraction also moving into a long-term soil carbon pool. The Amazon forest sink has therefore become increasingly skewed towards gains in the necromass pools, inducing a substantial lag in the probable atmospheric response. On the basis of the observed long-term increase in mortality rates, we estimate that the atmosphere has yet to see ~3.8 Pg of the Amazon necromass carbon produced since 1983 (see Methods), representing a 30% increase in necromass stocks. The modelled increase in Amazon necromass is twice the magnitude of the cumulative decadal decline in the live biomass sink from the 1990s to the 2000s (from 5.4 to 3.8 Pg C).

In summary, we find that the Amazon biomass carbon sink has started to decline, due to recent levelling of productivity increases, combined with a sustained long-term increase in tree mortality. This behaviour is at odds with expectations from models of a continually strong tropical

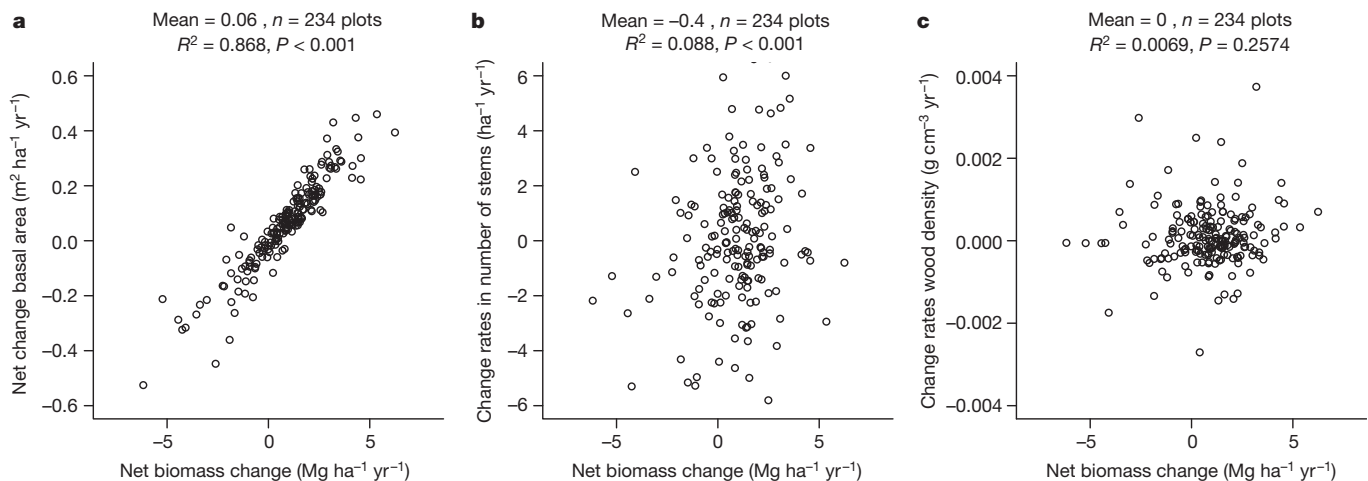


Figure 3 | Relationships between annual net change in biomass of individual plots and their annual change in basal area, stem numbers and wood density. a–c, The mean values of the rates of changes for basal area (a), stem numbers per hectare (b) and wood density (c) are given in each panel

along with the R^2 of the relationship with annual net biomass change and the P value of the linear relationship. The number of plots included is 234 (that is, those with data on change in basal area, stem numbers and wood density).

biomass sink⁶, and underlines how difficult it remains to predict the role of land-vegetation feedbacks in modulating global climate change^{7,10}. Investment in consistent, coordinated long-term monitoring on the ground is fundamental to determine the trajectory of the planet's most productive and diverse biome.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 9 April 2014; accepted 4 February 2015.

- Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P. & White, J. W. C. Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature* **488**, 70–72 (2012).
- Le Quéré, C. *et al.* The global carbon budget 1959–2011. *Earth System Science Data* **5**, 165–185 (2013).
- Pan, Y. *et al.* A large and persistent carbon sink in the world's forests. *Science* **333**, 988–993 (2011).
- Phillips, O. L. *et al.* Drought sensitivity of the Amazon rainforest. *Science* **323**, 1344–1347 (2009).
- Bugmann, H. & Bigler, C. Will the CO₂ fertilization effect in forests be offset by reduced tree longevity? *Oecologia* **165**, 533–544 (2011).
- Huntingford, C. *et al.* Simulated resilience of tropical rainforests to CO₂-induced climate change. *Nature Geosci.* **6**, 268–273 (2013).
- Booth, B. B. *et al.* High sensitivity of future global warming to land carbon cycle processes. *Environ. Res. Lett.* **7**, 024002 (2012).
- Feldpausch, T. R. *et al.* Tree height integrated into pantropical forest biomass estimates. *Biogeosciences* **9**, 3381–3403 (2012).
- Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. J. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**, 184–187 (2000).
- Friend, A. D. *et al.* Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO₂. *Proc. Natl Acad. Sci.* **111**, 3280–3285 (2013).
- Phillips, O. L. & Gentry, A. H. Increasing turnover through time in tropical forests. *Science* **263**, 954–958 (1994).
- Fisher, J. I., Hurr, G. C., Thomas, R. Q. & Chambers, J. Q. Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecol. Lett.* **11**, 554–563 (2008).
- Condit, R. Forest turnover, diversity, and CO₂. *Trends Ecol. Evol.* **12**, 249–250 (1997).
- Chambers, J. Q. *et al.* Response of tree biomass and wood litter to disturbance in a Central Amazon forest. *Oecologia* **141**, 596–611 (2004).
- van Breugel, M., Martínez-Ramos, M. & Bongers, F. Community dynamics during early secondary succession in Mexican tropical rain forests. *J. Trop. Ecol.* **22**, 663–674 (2006).
- Gatti, L. V. *et al.* Drought sensitivity of Amazonian carbon balance revealed by atmospheric measurements. *Nature* **506**, 76–80 (2014).
- Clark, D. A., Clark, D. B. & Oberbauer, S. F. Field-quantified responses of tropical rainforest aboveground productivity to increasing CO₂ and climatic stress, 1997–2009. *J. Geophys. Res.* **118**, 783–794 (2013).
- Wang, X. *et al.* A two-fold increase of carbon cycle sensitivity to tropical temperature variations. *Nature* **506**, 212–215 (2014).
- Marengo, J. A., Tomasella, J., Alves, L. M., Soares, W. R. & Rodriguez, D. A. The drought of 2010 in the context of historical droughts in the Amazon region. *Geophys. Res. Lett.* **38**, L12703 (2011).
- Jiménez-Muñoz, J. C., Sobrino, J. A., Mattar, C. & Malhi, Y. Spatial and temporal patterns of the recent warming of the Amazon forest. *J. Geophys. Res.* **118**, 5204–5215 (2013).
- da Costa, A. C. L. *et al.* Effect of 7 yr of experimental drought on vegetation dynamics and biomass storage of an eastern Amazonian rainforest. *New Phytol.* **187**, 579–591 (2010).
- Nepstad, D. C., Tohver, I. M., Ray, D., Moutinho, P. & Cardinot, G. Mortality of large trees and lianas following experimental drought in an Amazon forest. *Ecology* **88**, 2259–2269 (2007).
- Ryan, M. G., Phillips, N. & Bond, B. J. The hydraulic limitation hypothesis revisited. *Plant Cell Environ.* **29**, 367–381 (2006).
- Lieberman, D., Lieberman, M., Peralta, R. & Hartshorn, G. S. Mortality patterns and stand turnover rates in a wet tropical forest in Costa Rica. *J. Ecol.* **73**, 915–924 (1985).
- Thomas, S. C. in *Size- and Age-Related Changes in Tree Structure and Function* (eds Meinzer, F. C., Lachenbruch, B. & Dawson, T. E.) Ch. 2 33–64 (Springer, 2011).
- Bigler, C. & Veblen, T. T. Increased early growth rates decrease longevity of conifers in subalpine forests. *Oikos* **118**, 1130–1138 (2009).
- Di Filippo, A., Biondi, F., Maugeri, M., Schirone, B. & Piovesan, G. Bioclimate and growth history affect beech lifespan in the Italian Alps and Apennines. *Glob. Change Biol.* **18**, 960–972 (2012).
- LaDeau, S. L. & Clark, J. S. Rising CO₂ levels and the fecundity of forest trees. *Science* **292**, 95–98 (2001).
- Manusch, C., Bugmann, H., Heiri, C. & Wolf, A. Tree mortality in dynamic vegetation models – a key feature for accurately simulating forest properties. *Ecol. Modell.* **243**, 101–111 (2012).
- Saleska, S. R. *et al.* Carbon in Amazon forests: unexpected seasonal fluxes and disturbance-induced losses. *Science* **302**, 1554–1557 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements The RAINFOR forest monitoring network has been supported principally by the Natural Environment Research Council (grants NE/B503384/1, NE/D01025X/1, NE/I02982X/1, NE/F005806/1, NE/D005590/1 and NE/I028122/1), the Gordon and Betty Moore Foundation, and by the EU Seventh Framework Programme (GEOCARBON-283080 and AMAZALERT-282664). R.J.W.B. is funded by NERC Research Fellowship NE/I021160/1. O.P. is supported by an ERC Advanced Grant and is a Royal Society-Wolfson Research Merit Award holder. Additional data were supported by Investissement d'Avenir grants of the French ANR (CEBA: ANR-10-LABX-0025; TULIP: ANR-10-LABX-0041), and contributed by the Tropical Ecology Assessment and Monitoring (TEAM) Network, funded by Conservation International, the Missouri Botanical Garden, the Smithsonian Institution, the Wildlife Conservation Society and the Gordon and Betty Moore Foundation. This paper is 656 in the Technical Series of the Biological Dynamics of Forest Fragments Project (BDFFP-INPA/STRI). The field data summarized here involve vital contributions from many field assistants and rural communities in Bolivia, Brazil, Colombia, Ecuador, French Guiana, Guyana, Peru and Venezuela, most of whom have been specifically acknowledged elsewhere⁴. We additionally thank A. Alarcon, I. Amaral, P. P. Barbosa Camargo, I. F. Brown, L. Blanc, B. Burban, N. Cardozo, J. Engel, M. A. de Freitas, A. de Oliveira, T. S. Fredericksen, L. Ferreira, N. T. Hinojosa, E. Jiménez, E. Lenza, C. Mendoza, I. Mendoza Polo, A. Peña Cruz, M. C. Peñuela, P. Pétronelli, J. Singh, P. Maquirino, J. Serano, A. Sota, C. Oliveira dos Santos, J. Ybarnegaray and J. Ricardo for contributions. CNPq (Brazil), MCT (Brazil), Ministerio del Medio Ambiente, Vivienda y Desarrollo Territorial (Colombia), Ministerio de Ambiente (Ecuador), the Forestry Commission (Guyana), INRENA (Peru), SERNANP (Peru), and Ministerio del Ambiente para el Poder Popular (Venezuela) granted research permissions. We thank our deceased colleagues and friends, A. H. Gentry, J. P. Veillon, S. Almeida and S. Patiño for invaluable contributions to this work; their pioneering efforts to understand neotropical forests continue to inspire South American ecologists.

Author Contributions O.L.P., J.L. and Y.M. conceived the RAINFOR forest census plot network programme, E.G. and T.R.B. contributed to its development. R.J.W.B., O.L.P. and E.G. wrote the paper, R.J.W.B., O.L.P., T.R.F. and E.G. designed the study, R.J.W.B. carried out the data analysis, R.J.W.B., O.L.P., T.R.F., T.R.B., A.M.-M. and G.L.-G. coordinated data collection with the help of most co-authors, G.L.-G., O.L.P., S.L., T.R.B., T.R.F., R.J.W.B., J.T., E.G. and J.L. developed or contributed to analytical tools used in the analysis. All co-authors collected field data and commented on the manuscript.

Author Information Source data are available from http://dx.doi.org/10.5521/ForestPlots.net/2014_4. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.J.W.B. (r.brien@leeds.ac.uk).

R. J. W. Brien^{1*}, O. L. Phillips^{1*}, T. R. Feldpausch^{1,2}, E. Gloor¹, T. R. Baker¹, J. Lloyd^{3,4}, G. Lopez-Gonzalez¹, A. Monteagudo-Mendoza⁵, Y. Malhi⁶, S. L. Lewis^{1,7}, R. Vásquez Martínez⁵, M. Alexiades⁸, E. Álvarez Dávila⁹, P. Alvarez-Loayza¹⁰, A. Andrade¹¹, L. E. O. C. Aragão^{2,12}, A. Araujo-Murakami¹³, E. J. M. M. Arets¹⁴, L. Arroyo¹³, G. A. Aymard¹⁵, O. S. Bánki¹⁶, C. Baraloto^{17,18}, J. Barroso¹⁹, D. Bona²⁰, R. G. A. Boot²¹, J. L. C. Camargo¹¹, C. V. Castilho²², V. Chama²³, K. J. Chao^{1,24}, J. Chave²⁵, J. A. Comiskey²⁶, F. Cornejo Valverde²⁷, L. da Costa²⁸, E. A. de Oliveira²⁹, A. Di Fiore³⁰, T. L. Erwin³¹, S. Fauset¹, M. Forsthofer²⁹, D. R. Galbraith¹, E. S. Grahame¹, N. Groot¹, B. Hérault³², N. Higuchi¹¹, E. N. Honorio Coronado^{1,33}, H. Keeling¹, T. J. Killeen³⁴, W. F. Laurance³⁵, S. Laurance³⁵, J. Licona³⁶, W. E. Magnusson³⁷, B. S. Marimon²⁹, B. H. Marimon-Junior²⁹, C. Mendoza^{38,39}, D. A. Neill⁴⁰, E. M. Nogueira⁴¹, P. Núñez²³, N. C. Pallqui Camacho²³, A. Parada¹³, G. Pardo-Molina⁴², J. Peacock¹, M. Peña-Claros^{36,43}, G. C. Pickavance¹, N. C. A. Pitman^{10,44}, L. Poorter⁴³, A. Prieto⁴⁵, C. A. Quesada⁴¹, F. Ramírez⁴⁵, H. Ramírez-Angulo⁴⁶, Z. Restrepo⁹, A. Roopsind⁴⁷, A. Rudas⁴⁸, R. P. Salomão⁴⁹, M. Schwarz¹, N. Silva⁵⁰, J. E. Silva-Espejo²³, M. Silveira⁵¹, J. Stropp⁵², J. Talbot¹, H. ter Steege^{53,54}, J. Terán-Aguilar⁵⁵, J. Terborgh¹⁰, R. Thomas-Caesar⁵⁰, M. Toledo³⁶, M. Torello-Raventos^{56,57}, R. K. Umetzu²⁹, G. M. F. van der Heijden^{58,59,60}, P. van der Hout⁶¹, I. C. Guimarães Vieira⁴⁹, S. A. Vieira⁶², E. Vilanova⁴⁶, V. A. Vos^{42,63} & R. J. Zagt²¹

¹School of Geography, University of Leeds, Leeds LS2 9JT, UK. ²Geography, College of Life and Environmental Sciences, University of Exeter, Rennes Drive, Exeter EX4 4RJ, UK.

³Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK. ⁴School of Marine and Tropical Biology, James Cook University, Cairns, 4870 Queensland, Australia. ⁵Jardín Botánico de Missouri,

Prolongación Bolognesi Mze, Lote 6, Oxapampa, Pasco, Peru. ⁶Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford OX1 3QK, UK. ⁷Department of Geography, University College London, Pearson Building, Gower Street, London WC1E 6BT, UK. ⁸School of Anthropology and Conservation, Marlowe Building, University of Kent, Canterbury CT1 3EH, UK. ⁹Servicios Ecosistémicos y Cambio

Climático, Jardín Botánico de Medellín, Calle 73 no. 51 D-14, C.P. 050010, Medellín, Colombia. ¹⁰Center for Tropical Conservation, Duke University, Box 90381, Durham, North Carolina 27708, USA. ¹¹Biological Dynamics of Forest Fragment Project (INPA & STRI), C.P. 478, Manaus AM 69011-970, Brazil. ¹²National Institute for Space Research (INPE), Av. Dos Astronautas, 1758, São José dos Campos, São Paulo 12227-010, Brazil.

¹³Museo de Historia Natural Noel Kempf Mercado, Universidad Autónoma Gabriel Rene Moreno, Casilla 2489, Av. Irala 565, Santa Cruz, Bolivia. ¹⁴Alterra, Wageningen University and Research Centre, PO Box 47, 6700 AA Wageningen, The Netherlands.

¹⁵UNELLEZ-Guanare, Programa de Ciencias del Agro y el Mar, Herbario Universitario

(PORT), Mesa de Cavacas, Estado Portuguesa, 3350 Venezuela. ¹⁶Biodiversiteit en Ecosysteem Dynamica, University of Amsterdam, Postbus 94248, 1090 GE Amsterdam, The Netherlands. ¹⁷Institut National de la Recherche Agronomique, UMR EcoFoG, Campus Agronomique, 97310 Kourou, French Guiana. ¹⁸International Center for Tropical Botany, Department of Biological Sciences, Florida International University, Miami, Florida 33199, USA. ¹⁹Universidade Federal do Acre, Campus de Cruzeiro do Sul, Rio Branco, Brazil. ²⁰INRA, UMR 1137 "Ecologie et Ecophysiologie Forestiere" 54280 Champenoux, France. ²¹Tropenbos International, PO Box 232, 6700 AE Wageningen, The Netherlands. ²²Embrapa Roraima, Caixa Postal 133, Boa Vista, RR, CEP 69301-970, Brazil. ²³Universidad Nacional San Antonio Abad del Cusco, Av. de la Cultura N° 733, Cusco, Peru. ²⁴International Master Program of Agriculture, College of Agriculture and Natural Resources, National Chung Hsing University, Taichung 40227, Taiwan. ²⁵Université Paul Sabatier CNRS, UMR 5174 Evolution et Diversité Biologique, Bâtiment 4R1, 31062 Toulouse, France. ²⁶Northeast Region Inventory and Monitoring Program, National Park Service, 120 Chatham Lane, Fredericksburg, Virginia 22405, USA. ²⁷Andes to Amazon Biodiversity Program, Puerto Maldonado, Madre de Dios, Peru. ²⁸Universidade Federal do Para, Centro de Geociencias, Belem, CEP 66017-970 Para, Brazil. ²⁹Universidade do Estado de Mato Grosso, Campus de Nova Xavantina, Caixa Postal 08, CEP 78.690-000, Nova Xavantina MT, Brazil. ³⁰Department of Anthropology, University of Texas at Austin, SAC Room 5.150, 2201 Speedway Stop C3200, Austin, Texas 78712, USA. ³¹Department of Entomology, Smithsonian Institution, PO Box 37012, MRC 187, Washington DC 20013-7012, USA. ³²Cirad, UMR Ecologie des Forêts de Guyane, Campus Agronomique, 97310 Kourou, French Guiana. ³³Instituto de Investigaciones de la Amazonía Peruana, Av. A. José Quiñones km 2.5, Iquitos, Peru. ³⁴World Wildlife Fund, 1250 24th Street NW, Washington DC 20037, USA. ³⁵Centre for Tropical Environmental and Sustainability Science (TESS) and School of Marine and Environmental Sciences, James Cook University, Cairns, Queensland 4878, Australia. ³⁶Instituto Boliviano de Investigación Forestal, C.P. 6201, Santa Cruz de la Sierra, Bolivia. ³⁷National Institute for Research in Amazonia (INPA), C.P. 478, Manaus, Amazonas, CEP 69011-970, Brazil. ³⁸FOMABO, Manejo Forestal en las Tierras Tropicales de Bolivia, Sacta, Bolivia. ³⁹Escuela de Ciencias Forestales (ESFOR), Universidad Mayor de San Simón (UMSS), Sacta, Bolivia. ⁴⁰Universidad Estatal Amazónica, Facultad de Ingeniería Ambiental, Paso lateral km 2 1/2 via Napo, Puyo, Pastaza, Ecuador. ⁴¹National Institute for Research in Amazonia (INPA), C.P. 2223, 69080-971, Manaus, Amazonas, Brazil. ⁴²Universidad Autonoma del Beni, Campus Universitario, Av. Ejército Nacional, Riberalta, Beni, Bolivia. ⁴³Forest Ecology and Forest Management Group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands. ⁴⁴The Field Museum, 1400 South Lake Shore Drive, Chicago, Illinois 60605-2496, USA. ⁴⁵Universidad Nacional de la Amazonía Peruana, Iquitos, Loreto, Peru. ⁴⁶Instituto de Investigaciones para el Desarrollo Forestal (INDEFOR), Universidad de Los Andes, Facultad de Ciencias Forestales y Ambientales, Conjunto Forestal, C.P. 5101, Mérida, Venezuela. ⁴⁷Iwokrama International Centre for Rainforest Conservation and Development, 77 High Street Kingston, Georgetown, Guyana. ⁴⁸Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Ciudad Universitaria, Carrera 30 No 45-03, Edificio 425, C.P. 111321, Bogotá, Colombia. ⁴⁹Museu Paraense Emilio Goeldi, Av. Magalhães Barata, 376 - São Braz, CEP 66040-170, Belém PA, Brazil. ⁵⁰UFRA, Av. Presidente Tancredo Neves 2501, CEP 66.077-901, Belém, Pará, Brazil. ⁵¹Museu Universitário, Universidade Federal do Acre, Rio Branco AC 69910-900, Brazil. ⁵²European Commission – DG Joint Research Centre, Institute for Environment and Sustainability, Via Enrico Fermi 274, 21010 Ispra, Italy. ⁵³Naturalis Biodiversity Center, PO Box, 2300 RA, Leiden, The Netherlands. ⁵⁴Ecology and Biodiversity Group, Utrecht University, PO Box 80084, 3508 TB Utrecht, The Netherlands. ⁵⁵Museo de Historia Natural Alcide D'Orbigny, Av. Potosi no 1458, Cochabamba, Bolivia. ⁵⁶School of Earth and Environmental Science, James Cook University, Cairns, Queensland 4870, Australia. ⁵⁷Centre for Tropical Environmental and Sustainability Science (TESS) and School of Marine and Tropical Biology, James Cook University, Cairns, Queensland 4878, Australia. ⁵⁸Northumbria University, School of Geography, Ellison Place, Newcastle upon Tyne, Newcastle NE1 8ST, UK. ⁵⁹University of Wisconsin, Milwaukee, Wisconsin 53202, USA. ⁶⁰Smithsonian Tropical Research Institute, Apartado Postal 0843-03092, Panamá, Republic of Panama. ⁶¹Van der Hout Forestry Consulting, Jan Trooststraat 6, 3078 HP Rotterdam, The Netherlands. ⁶²Universidade Estadual de Campinas, NEPAM, Rua dos Flamboyants, 155- Cidade Universitária Zeferino Vaz, Campinas, CEP 13083-867, Sao Paulo, Brazil. ⁶³Centro de Investigación y Promoción del Campesinado, regional Norte Amazónico, C/ Nicanor Gonzalo Salvatierra N° 362, Casilla 16, Riberalta, Bolivia.

*These authors contributed equally to this work.

METHODS

Forest biometric data. Mature forests were sampled throughout the forested lowland tropical areas of South America (below 1,500 m above sea level) that receive at least 1,000 mm of rainfall annually. To be included in this study, permanent sample plots were required to have two or more censuses. Immature or open forests, and those known to have had anthropogenic disturbances owing to fire or selective logging, were excluded. The plots are geographically well dispersed throughout the Amazon Basin (Extended Data Fig. 1), covering every tropical South American country except Suriname. Supplementary Table 1 includes a complete list of plots included in this study with the respective size, start and end date for the censuses included in this analysis, and names of the main researchers for each plot. A full manual for plot establishment and tree measurements of the RAINFOR plot network can be found in ref. 31.

Of the total 321 plots, 232 are from the RAINFOR network. In addition, we compiled biomass dynamics data for 89 plots from published studies, mostly from one site for 2001 to 2003 (DUK) (see Supplementary Table 1). For these plots, we simply used the available biomass data as published. Note that these studies do not apply the same allometric equations, and may have slightly different measurements protocols and census interval corrections. While as a general rule all trees with stem diameters greater than 100 mm were included in this analysis, palms (Arecaceae) or coarse herbs of the genus *Phenakospermum* were excluded for a few plots (19) due to changes in measurement protocols over time in these plots. In addition, for a few plots only trees ≥ 130 or ≥ 200 mm in diameter were recorded in the first census(es). In these cases, we either standardized the biomass data in the first census(es) to trees ≥ 100 mm using the ratio of biomass for trees ≥ 100 and ≥ 200 mm (seven plots) of later censuses, or we used the slightly different minimum size threshold for the full period, including only trees ≥ 130 mm (for two plots). For full details on these specific issues see the online source data. For analysis purposes, plots smaller than 0.5 ha that were within 1 km or less of one another were merged, to give a total of 274 'sample units'. The mean size across all sample units was 1.24 ha, and the mean total monitoring period was 11.1 years. In total, the study monitored 343 ha for a combined total of 4,620 ha years, involving more than 850,000 tree measurements on around 189,000 individual trees larger than 10 cm diameter.

The standard protocol for tree measurements in the field is to measure diameter at breast height, defined as 1.3 m from the base of the stem. For non-cylindrical stems owing to buttresses or other deformities the point of measurement is raised approximately 50 cm above the deformity. The exact height of the point of measurement (POM) was recorded and marked on the trees to ensure that future measurements were taken at the same point. For those trees where buttress growth threatened to reach the initial POM, we raised the height of diameter measurement to a new POM, located sufficiently high above the buttresses to avoid interference of buttresses with diameter measurements at subsequent censuses. If a change in POM was made, we recorded both the diameter at the original POM and the new POM, thus creating two disjoint series of diameters measured at different heights. To avoid potential biases that can result from not accounting for the POM movement, following ref. 32 we computed a new diameter series that was calculated as the mean of: (1) diameter measurements standardized to the new (final) POM, obtained by multiplication of measurements at the original POM by the ratio between diameter measurements at the new and original POM, and (2) diameter measurements standardized to the original POM, by multiplying measurements at the new POM by the ratio between diameter measurements at the original and new POM. The outcome of our analysis was robust with respect to the method of dealing with POM changes, giving similar results using several alternative approaches for dealing with POM changes including the technique described previously¹⁷ in which diameter gains at the new POM are added to the diameter at the original POM. Following ref. 32 we used several techniques to avoid or minimise potential errors arising from missing diameter values, typographical errors, or extreme diameter growth ≥ 4 cm yr⁻¹ or total diameter growth ≤ -0.5 cm across a single census interval (that is, losing 0.5 cm, as trees may shrink by a small amount due to hydrostatic effects in times of drought, and measurement errors can be both positive and negative). For stems belonging to species known to experience very high growth rates or noted as having damaged stems we accepted these values. We used interpolation, where possible, else extrapolation to correct errors. If neither of these procedures were possible we used the mean growth rate of all dicotyledonous stems in the same plot census, belonging to the same size class, with size classes defined as $10 \leq \text{diameter} < 20$ cm, $20 \leq \text{diameter} < 40$ cm, and $\text{diameter} \geq 40$ cm, to estimate the missing diameter value. Of all stem growth increments, for 1.7% per census we assigned interpolated estimates of diameter, for 0.9% we used extrapolated estimates, and for 1.5% we used mean growth rates.

Computing above ground biomass, sampling effects and scaling up sink estimates. We converted diameter measurements to AGB estimates using allometric equations described previously⁸, which include terms for wood density, diameter and tree height. Tree height was estimated based on established diameter-height

relations that vary between the different regions of Amazonia⁸. Wood density values were extracted from a global wood density database (<http://datadryad.org/handle/10255/dryad.235>; ref. 33). In cases where a stem was unidentified or where no taxon-specific wood density data were available, we applied the appropriate genus or family-specific wood density values. If none of those was available, the mean wood density of all identified dicotyledonous tree stems in the plot was applied. In our analysis 80% of the trees were identified to species level, 94% to genus level, and 97% to family level. All data on tree diameter, taxonomy, and associated botanical vouchers are curated under the <https://www.forestplots.net/> web application and database³⁴.

The magnitude of the biomass sink for the forested area of the Amazon Basin for the 1990s and 2000s was estimated by multiplying the magnitude of total biomass change with an estimated area of intact forest, including all open and closed, evergreen and deciduous forests for tropical South America (6.29×10^8 ha, according to Global Land Cover map 2000; ref. 35). For this calculation we also included biomass components that were not directly measured, assuming that these pools responded proportionally to the measured above ground biomass in trees bigger than 10 cm in diameter. It has been shown using destructive measurements of stand biomass in central Amazonia that lianas and trees smaller than 100 mm in diameter represent an additional fraction of $\sim 9.9\%$ the measured AGB (in trees ≥ 10 cm in diameter³⁶), and below ground biomass a fraction of $\sim 37\%$ the AGB³⁶. We assumed that 50% of biomass is carbon³⁷.

Analysing time trends and statistical analysis. The longer a census interval, the greater the proportion of growth that cannot be directly observed within the interval, due to the growth of initially recorded trees that subsequently die during the interval, and the growth of unrecorded trees that both recruit and die during the interval^{38–40}. Hence, variation in census interval lengths in plots over time will affect estimates of woody productivity and mortality rates⁴⁰, potentially biasing the long-term trends if not accounted for. Using established procedures³², we therefore explicitly corrected for the influence of varying census interval length, by estimating the following two unobserved components: (1) unobserved recruits, that is, the cohort of recruits that both enter and die between two successive censuses, and (2) unobserved biomass growth and mortality, due to the growth of trees after the final census that a tree was recorded alive. To correct for unobserved recruits, we first estimated the number of unobserved recruits (U_t) as the number of stems in the plot (N) multiplied by the annual recruitment rate (R) multiplied by the mean annual mortality rate (M) multiplied by the census interval length (t): $U_t = N \times R \times M \times t$. We assumed that the diameter of these trees was 100 mm plus growth for one-third of the interval using the median growth rate for trees in the 100–200 mm size class. The biomass of each tree was estimated by applying the regionally appropriate allometric equation⁸, using the plot mean wood density. To correct for unobserved growth and mortality due to trees dying within an interval, we assumed that all trees that died during the interval to have died at the mid-point, and assigned growth up to this mid-point, estimated as the median growth of all trees in the plot within the same size class. Full details of the procedure have been described previously³². These estimates of the unobserved biomass dynamics usually accounted for only a small proportion of the total woody productivity and mortality (respectively 2.28% and 2.74%, on average).

Mean time trends of biomass dynamics (black lines in Fig. 1 and Extended Data Figs 3 and 5–7) were calculated for each month since 1983 as the weighted mean across all sample units. As plots vary in total area monitored, we used an empirical weighting procedure to account for differences between plots in sampling effort by weighting according to the square root of plot area^{4,41}. Confidence intervals (95%) were estimated using weighted bootstrap sampling.

To estimate long-term trends in biomass dynamics (cf. Figure 1), we first used general additive mixed models (GAMM) from the *gamm4* R package⁴². Estimates of the long-term trends were performed by regressing the mid-point of each census interval (Extended Data Fig. 2) against the rate of change (net change, mortality or gains). Here, systematic plot effects were explicitly accounted for by using plot as a random effect in the model. This avoided switches over time in the exact set of plots being monitored influencing the long-term trends. As census interval length and plot sizes varied, we weighted each data point in the regression by the product of the census interval length (in years) times the square root of plot size (in hectares), as suggested previously⁴¹. We estimated the linear slope of the long-term trend using the *lme4* package⁴³. In an identical way to the GAMM, we accounted for plot effects and added weights to the regression. To test whether the estimated time trends were robust to different plots being sampled over different timeframes, we also repeated the above analysis over shorter time windows (1990–2011.5, 1995–2011.5 and 2000–2011.5) keeping the set of plots used completely constant. Results of this analysis are shown in Extended Data Fig. 3.

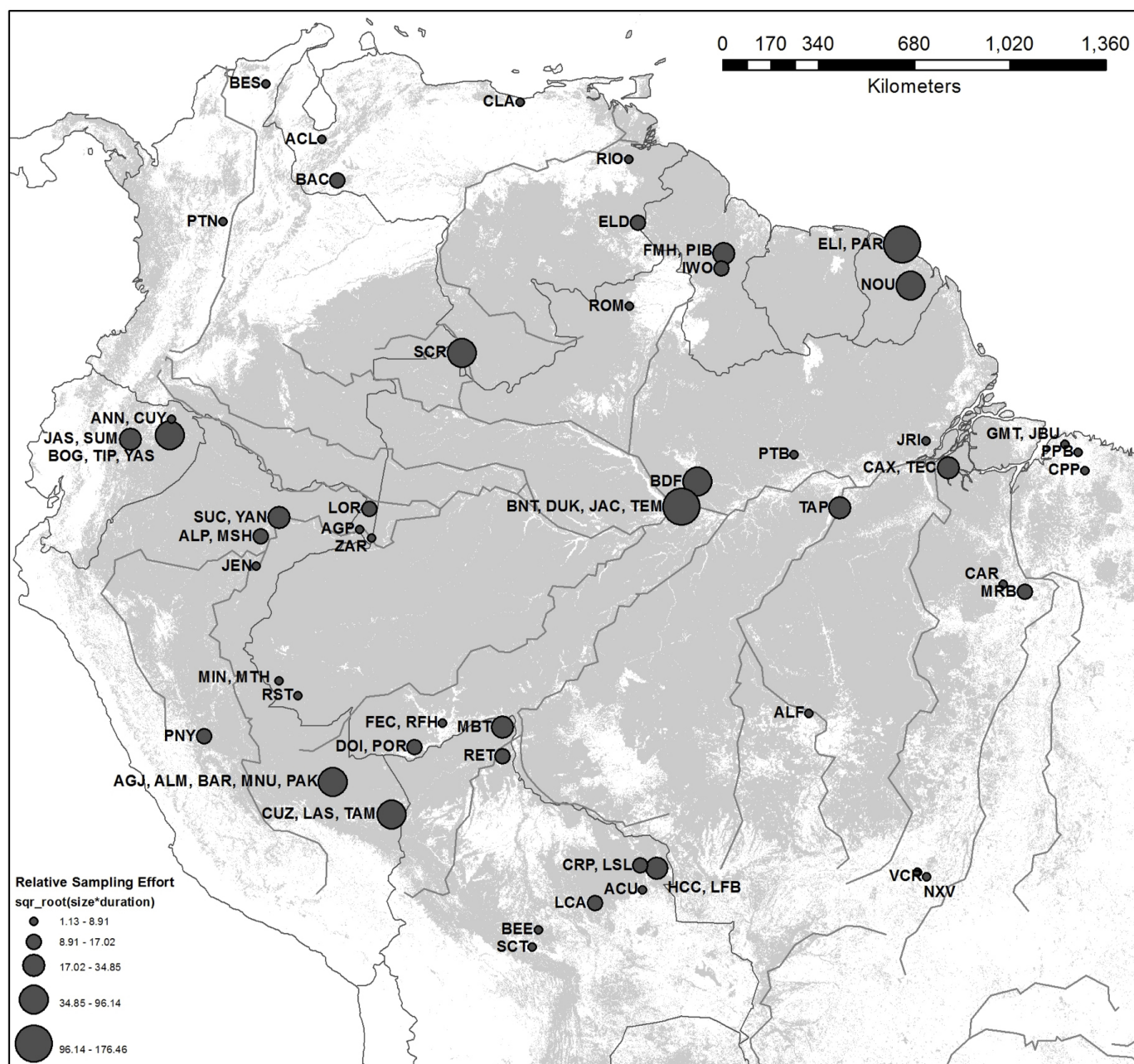
The approaches using GAMM and the linear slope calculations are parametric and assume normally distributed data, while census-level data on AGB mortality and net AGB change are non-normally distributed, showing respectively right-skewed

and left-skewed distributions. Thus the observed time series for AGB mortality and net AGB change do not strictly meet the criteria for this type of parametric analysis, although it might be expected from the central limit theorem that with sufficiently large data sets the regression analyses would still have validity. To test explicitly the robustness of our estimates for the models of net change and mortality with regard to violation of the normality assumption for ordinary least squares analysis, we used a rank-based estimator for linear models available from the Rfit-package⁴⁴. This shows that slopes for AGB net change and mortality are similar or else of larger magnitude using non-parametric tests (that is, slope net change = $-0.057 \text{ Mg ha}^{-1} \text{ yr}^{-2}$, $P < 0.001$, slope mortality = $0.061 \text{ Mg ha}^{-1} \text{ yr}^{-2}$, $P < 0.001$, compared respectively to values of $-0.034 \text{ Mg ha}^{-1} \text{ yr}^{-2}$ and $0.051 \text{ Mg ha}^{-1} \text{ yr}^{-2}$ using the parametric techniques). A test of non-parametric rank based estimations of the slopes of the change in standing biomass, or mortality on a per stem basis (Extended Data Fig. 6), and of changes in stem numbers and number of trees dying and recruiting per hectare (Extended Data Fig. 7), or basal area changes (Extended Data Fig. 8), show similar results to that of the parametric tests: there is a significant decrease in net change of standing biomass per stem ($P = 0.0014$), no trend in the losses on a per-stem basis ($P = 0.47$), a significant decrease in the change in the number of stems per hectare ($P < 0.001$), marginally significant increase in number of recruits ($P = 0.051$), a significant increase in the number of trees dying ($P < 0.001$), a significant decrease in net basal area change ($P < 0.001$), and a significant increase in basal area mortality ($P < 0.001$).

A second method for calculating the long-term trends in biomass dynamics involved estimating the slopes of the time trends for individual plots (Fig. 2). We did this only for those plots that had at least three census intervals, and more than 10 years of total monitoring length with at least one census interval in the 1990s and in the 2000s. These stricter selection criteria were designed to allow us to focus on a core set of data most likely to capture long-term patterns in regional biomass dynamics. Slopes of biomass dynamics metrics were seldom statistically significant ($P = 0.95$) within plots, due to the stochastic nature of the dynamics data (Extended Data Fig. 2). We calculated the mean of the slopes across all plots weighted by the product of square root of plot area times the total census interval length. A *t*-test was used to test whether the mean values were significantly different from zero.

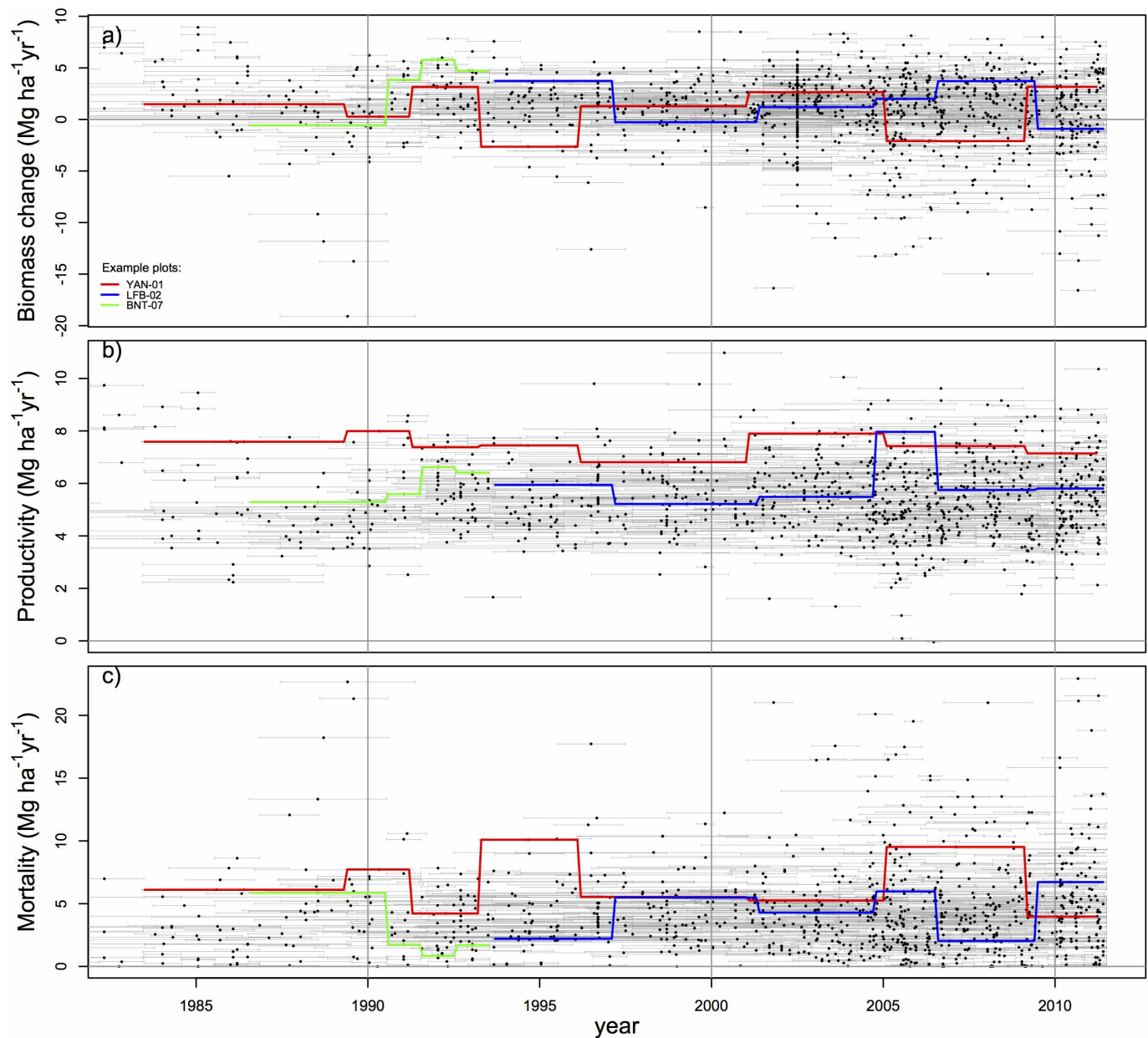
All analyses were performed using the R statistical platform, version 3.0.2 (ref. 45). No statistical methods were used to predetermine sample size.

31. Phillips, O., Baker, T., Brien, R. & Feldpausch, T. RAINFOR field manual for plot establishment and remeasurement. http://www.rainfor.org/upload/ManualsEnglish/RAINFOR_field_manual_version_June_2009_ENG.pdf (2010).
32. Talbot, J. *et al.* Methods to estimate aboveground wood productivity from long-term forest inventory plots. *For. Ecol. Management* **320**, 30–38 (2014).
33. Chave, J. *et al.* Towards a worldwide wood economics spectrum. *Ecol. Lett.* **12**, 351–366 (2009).
34. Lopez-Gonzalez, G., Lewis, S. L., Burkitt, M. & Phillips, O. L. ForestPlots.net: a web application and research tool to manage and analyse tropical forest plot data. *J. Veg. Sci.* **22**, 610–613 (2011).
35. Bartholomé, E. & Belward, A. GLC2000: a new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **26**, 1959–1977 (2005).
36. Phillips, O. L., Lewis, S. L., Baker, T. R., Chao, K. J. & Higuchi, N. The changing Amazon forest. *Phil. Trans. R. Soc. Lond. B* **363**, 1819–1827 (2008).
37. Chave, J. *et al.* Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* **145**, 87–99 (2005).
38. Sheil, D. & May, R. M. Mortality and recruitment rate evaluations in heterogeneous tropical forests. *J. Ecol.* **84**, 91–100 (1996).
39. Malhi, Y. *et al.* The above-ground coarse wood productivity of 104 Neotropical forest plots. *Glob. Change Biol.* **10**, 563–591 (2004).
40. Lewis, S. L. *et al.* Tropical forest tree mortality, recruitment and turnover rates: calculation, interpretation and comparison when census intervals vary. *J. Ecol.* **92**, 929–944 (2004).
41. Muller-Landau, H. C., Detto, M., Chisholm, R. A., Hubbell, S. P. & Condit, R. in *Forests and Global Change ecological reviews* (eds Coomes, D., Burslem, D. F. R. P. & Simonson, W. D.) Ch. 14 462 (Cambridge Univ. Press, 2014).
42. Wood, S. gamm4: Generalized additive mixed models using mgcv and lme4. R package version 0.1–2. Available at <http://www.inside-r.org/packages/gamm4/versions/0-1-2> (2011).
43. Bates, D., Maechler, M., Bolker, B. & Walker, S. lme4: Linear mixed-effects models using Eigen and S4. R package version, 1.0–4. Available at <http://www.inside-r.org/packages/lme4/versions/1-0-4> (2013).
44. Klok, J. D. & McKean, J. W. Rfit: Rank-based estimation for linear models. *Rem. J.* **4**, 57–64 (2012).
45. R Development Core Team. R: A Language and Environment for Statistical Computing. Available at <http://www.R-project.org/> (2013).



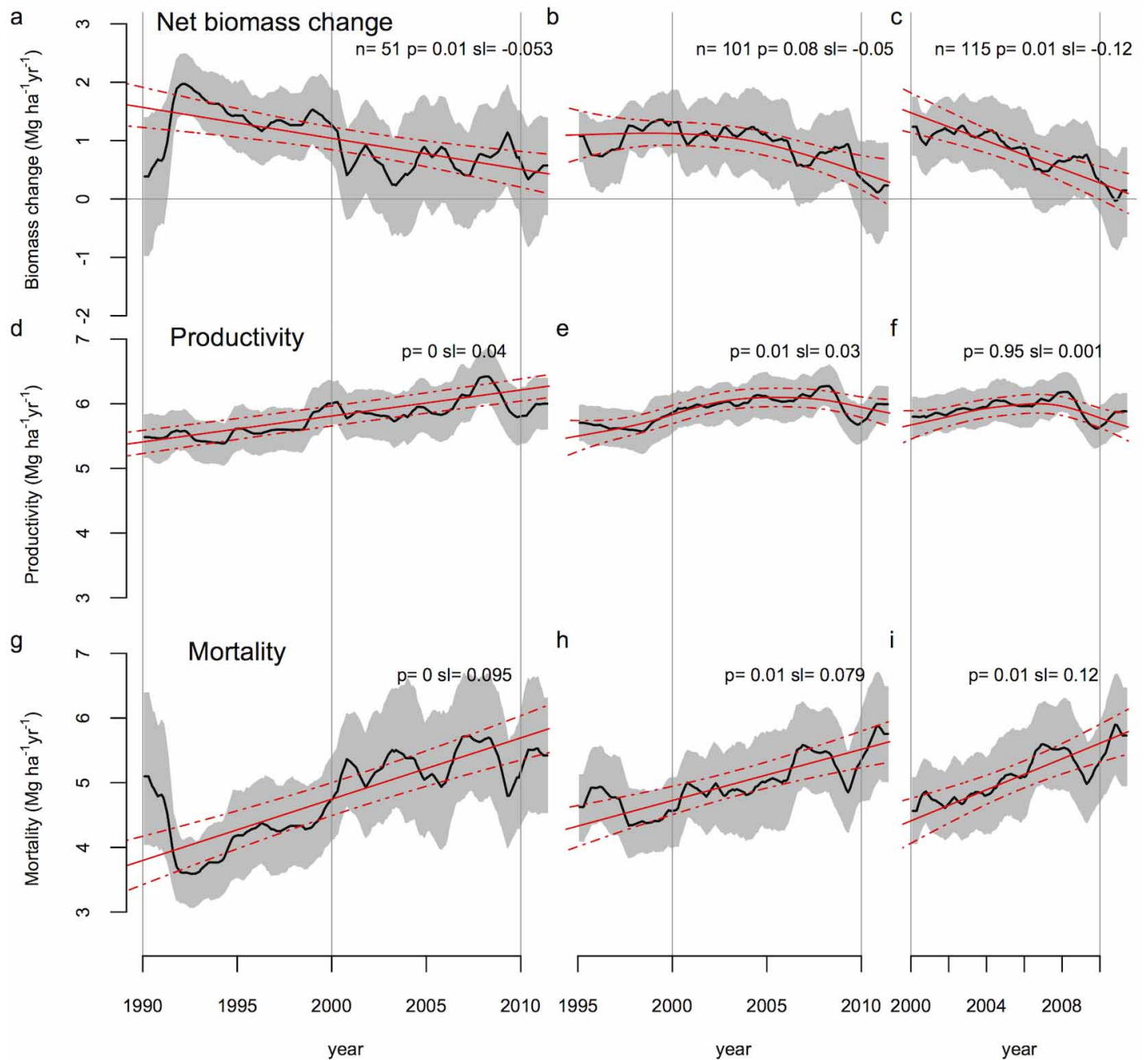
Extended Data Figure 1 | Map showing locations of plots included in this study. The three-letter codes refer to plot codes (see Supplementary Table 1). Adjacent plots (<50 km apart) are shown as one for display purposes. Size of the dots corresponds to the relative sampling effort at that location which is

calculated as the square root of plot size multiplied by square root of census length. The grey area shows the cover of all open and closed, evergreen and deciduous forests for tropical South America, according to Global Land Cover map 2000 (ref. 35).



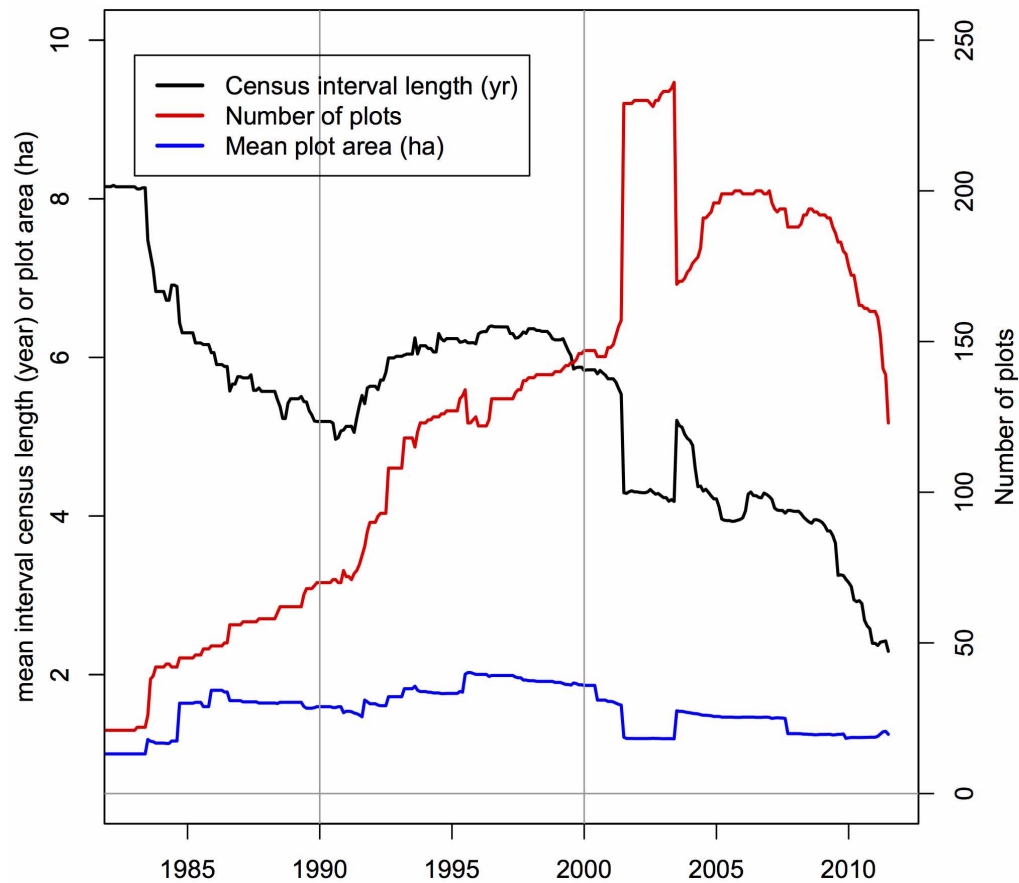
Extended Data Figure 2 | Scatterplot of mid-interval date against net AGB change, AGB productivity and AGB loss due to mortality for all data points and plots used in this analysis. a, Biomass change. b, Productivity. c, Mortality. Points indicate the mid-census interval date, while horizontal

error-bars connect the start and end date for each census interval. To illustrate variation in net AGB change over time within individual plots, examples of time series for three individual plots are shown as lines.



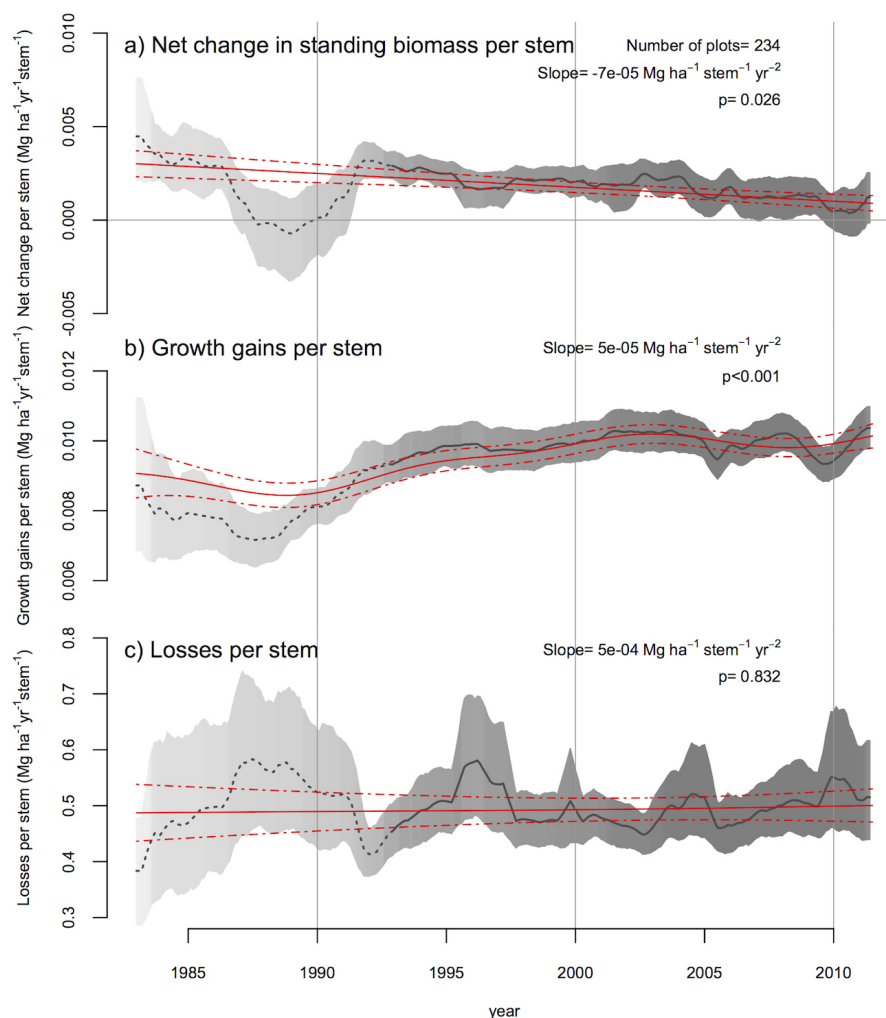
Extended Data Figure 3 | Time trends of subsets of net above-ground biomass change, above-ground woody productivity and mortality rates for plots that were continuously monitored throughout, for the periods 1990–2011, 1995–2011 and 2000–2011. Locations for the set of plots included in the analysis for the different periods are shown in the maps in lower panels.

The red lines indicate the best model fit for the long-term trends using General Additive Mixed Models (GAMM) accounting explicitly for differences in dynamics between plots (red lines denote overall mean, broken lines denote s.e.m.). Estimated long-term (linear) mean slopes (sl), P values and sample sizes (n) are indicated (see Methods).



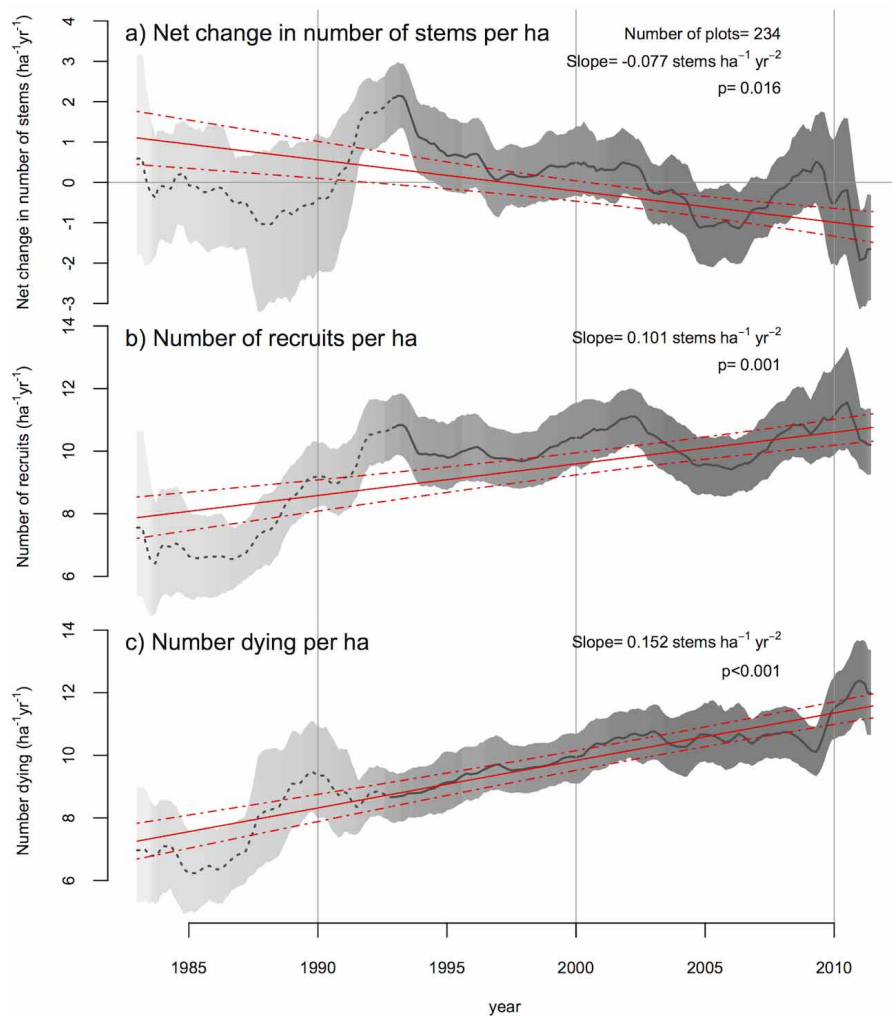
Extended Data Figure 4 | Mean number of plots, interval census length and area of all plots. The mean number of plots (red lines), mean interval census length (black lines) and mean plot area (blue lines) are shown. Note that

the increased sampling in 2002 to 2004 is largely due to the short-term addition of 72 plots from one site (Ducke, north of Manaus), but this has no discernible effect on averaged biomass dynamics (Fig. 1).



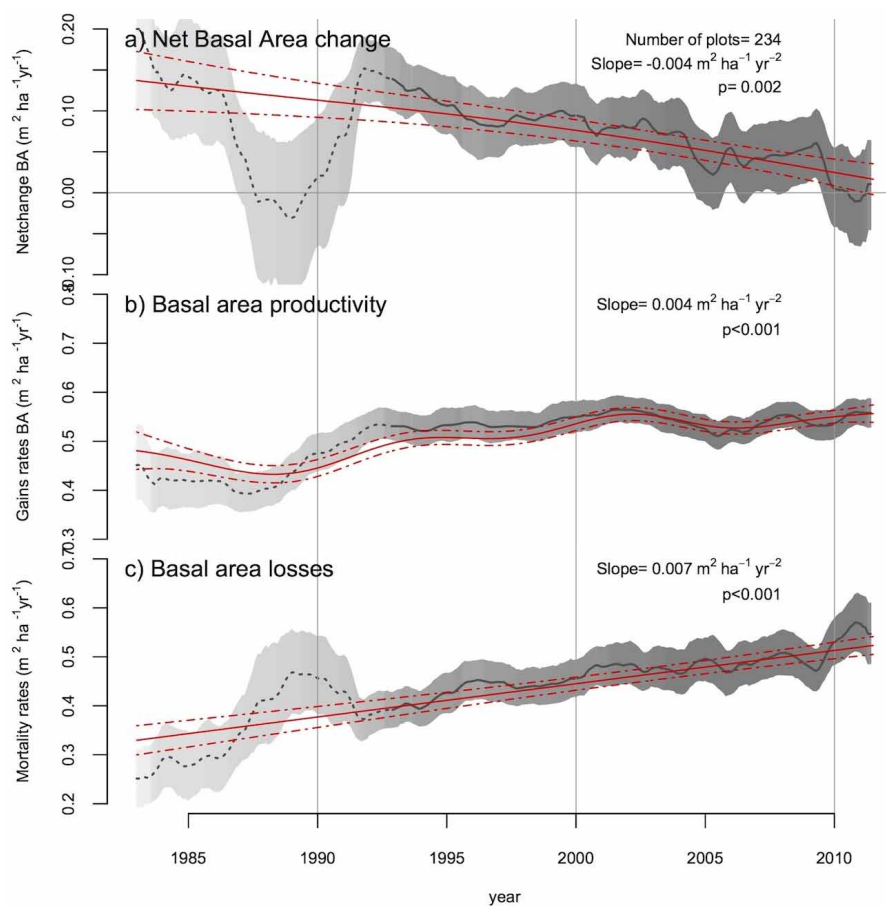
Extended Data Figure 5 | Biomass change, growth gains and mortalities on a per live stem basis. **a**, Mean net biomass change on a per live stem basis (that is, net biomass change per stem). **b**, Mean growth gains per live tree (that is, mean biomass accumulation of individual trees). **c**, Mortality losses per stem. Analyses are based on 234 plots, excluding published studies without

available stem-by-stem data. The red lines indicate the best model fit for the long-term trends using General Additive Mixed Models (GAMM) accounting explicitly for differences in dynamics between plots (red lines denote overall mean, broken lines denote s.e.m.). Estimated long-term (linear) mean slopes and significance levels are indicated (see Methods).



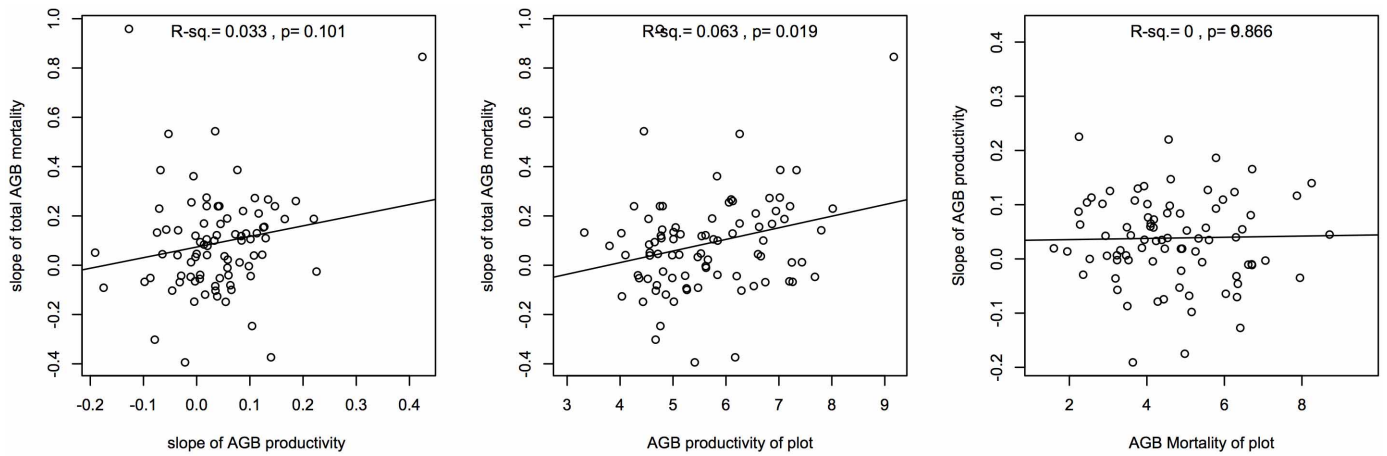
Extended Data Figure 6 | Rates of change in number of stems plus annualized fluxes of stems bigger than 10 cm in diameter. a–c, Mean net change in number of stems (a), number of recruits (b), and number of dying trees (c). Analyses are based on 234 plots, excluding published studies without available stem-by-stem data. The red lines indicate the best model fit

for the long-term trends using General Additive Mixed Models (GAMM) accounting explicitly for differences in dynamics between plots (red lines denote overall mean, broken lines denote s.e.m.). Estimated long-term (linear) mean slopes and significance levels are indicated (see Methods).



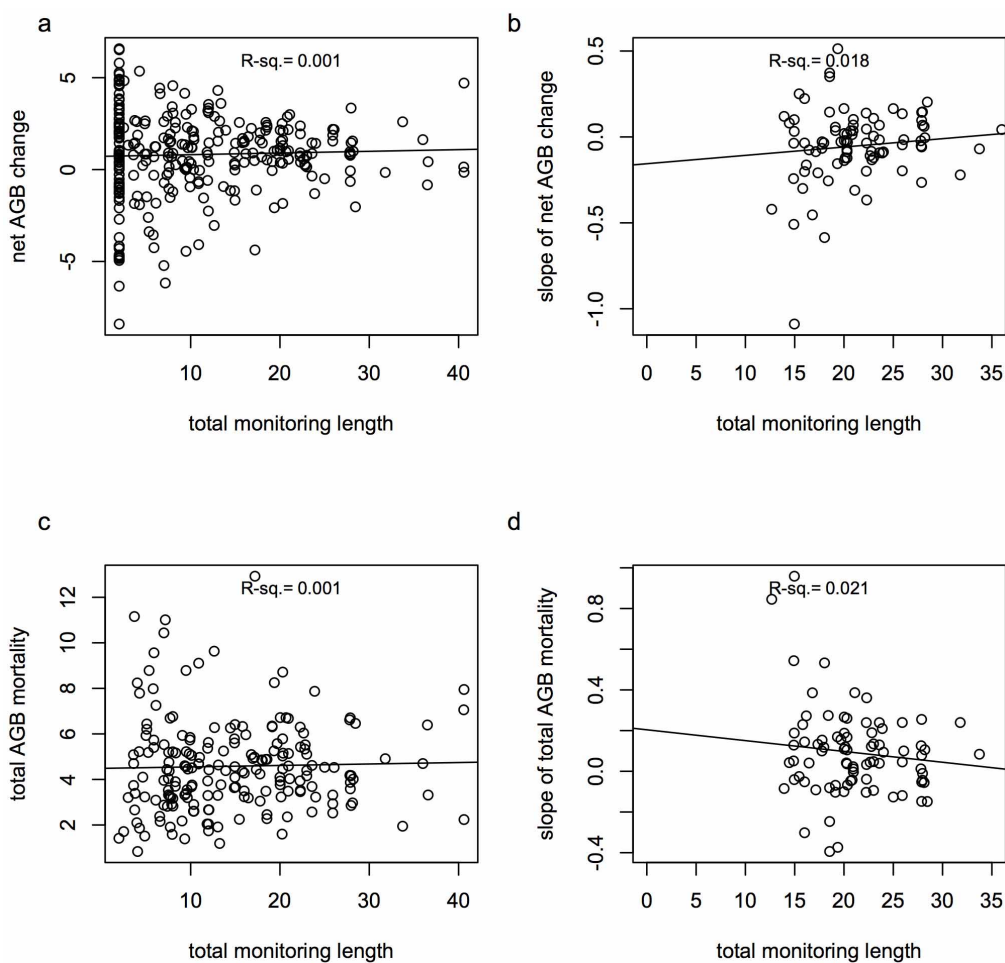
Extended Data Figure 7 | Basal area change, productivity and mortality.
a, Mean net basal area change. **b**, Mean basal area productivity. **c**, Mean basal area mortality. Analyses are based on 234 plots, excluding published studies without available basal-area data. The red lines indicate the best model fit for

the long-term trends using General Additive Mixed Models (GAMM) accounting explicitly for differences in dynamics between plots (red lines denote overall mean, broken lines denote s.e.m.). Estimated long-term (linear) mean slopes and significance levels are indicated (see Methods).



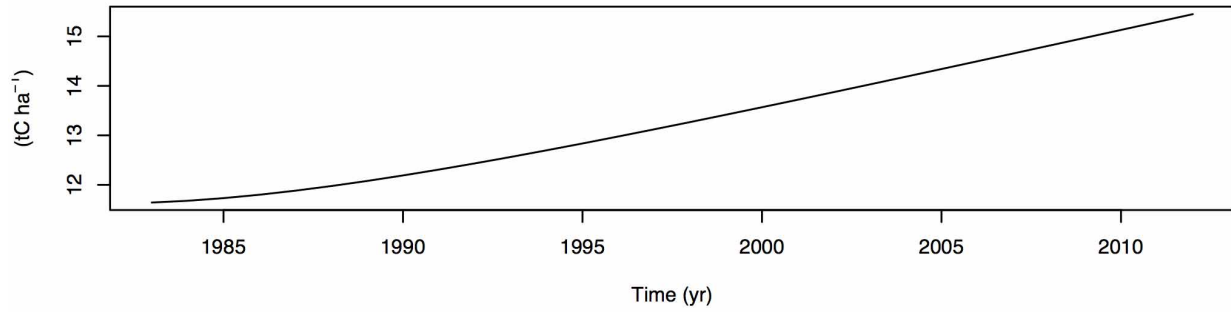
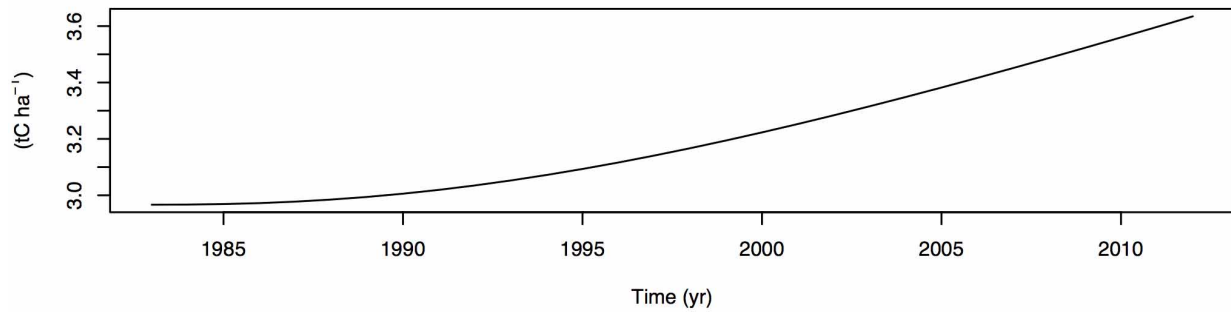
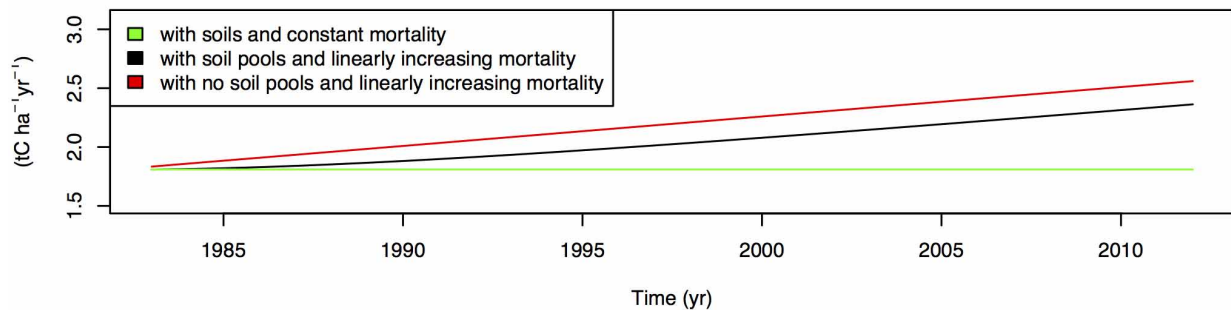
Extended Data Figure 8 | Relationship among plots between mean and slopes of AGB mortality and AGB productivity. **a**, Scatterplot of the slope of AGB mortality of individual plots against the slope of AGB productivity of plots. **b**, Scatterplot of the slope of AGB loss due to mortality of individual plots against the mean AGB productivity of plots. **c**, Scatterplot of

the slope of AGB productivity of individual plots against the mean AGB loss due to mortality of plots. The set of plots used in this analysis (117 plots, 87 units) includes only those that had at least 10 years of data and at least three census intervals (that is, same criteria as plots shown in Fig. 2).



Extended Data Figure 9 | Net AGB change or loss due to mortality versus the total monitoring length of plots, and the slope of net AGB change or mortality versus the total monitoring length of plots. a–d, Scatterplots of net AGB change (a) or net AGB loss due to mortality of individual plots (c) against the total monitoring length of plots, and the slope of net AGB change (b) or slope of AGB mortality of individual plots (d) against the total

monitoring length of plots. None of the relationships are significant ($P > 0.05$). Note that the plots (117 plots, 87 units) used in b and d are only those that had at least 10 years of data and at least three census intervals (that is, same criteria as plots shown in Fig. 2). See Supplementary Information for discussion of these results.

a) Necromass for linearly increasing mortality**b) Soil organic matter for linearly increasing mortality****c) Flux to atmosphere from necromass**

Extended Data Figure 10 | Modelled estimates of the effects of linearly increasing mortality on necromass stocks and soil organic-matter stocks. **a**, Necromass stocks. **b**, Soil organic matter stocks. **c**, The estimated fluxes of carbon from the forest to the atmosphere in three scenarios: (1) assuming constant mortality rate and a lag in decomposition of dead-tree biomass

(green), (2) assuming an increasing mortality rate similar to the observed trend (Fig. 1c) and a lag in decomposition as modelled (black), and (3) with increasing mortality but with all dead-tree biomass instantly respired (red). See Supplementary Information for discussion of these results.

Polyploidy can drive rapid adaptation in yeast

Anna M. Selmecki^{1,2,3†}, Yosef E. Maruvka^{4,5}, Phillip A. Richmond^{6,7}, Marie Guillet^{1,2,3}, Noam Shores⁸, Amber L. Sorenson^{6,7}, Subhajyoti De^{9,10,11}, Roy Kishony^{12,13}, Franziska Michor^{4,5}, Robin Dowell^{6,7} & David Pellman^{1,2,3,14}

Polyploidy is observed across the tree of life, yet its influence on evolution remains incompletely understood^{1–4}. Polyploidy, usually whole-genome duplication, is proposed to alter the rate of evolutionary adaptation. This could occur through complex effects on the frequency or fitness of beneficial mutations^{2,5–7}. For example, in diverse cell types and organisms, immediately after a whole-genome duplication, newly formed polyploids missegregate chromosomes and undergo genetic instability^{8–13}. The instability following whole-genome duplications is thought to provide adaptive mutations in microorganisms^{13,14} and can promote tumorigenesis in mammalian cells^{11,15}. Polyploidy may also affect adaptation independently of beneficial mutations through ploidy-specific changes in cell physiology¹⁶. Here we perform *in vitro* evolution experiments to test directly whether polyploidy can accelerate evolutionary adaptation. Compared with haploids and diploids, tetraploids undergo significantly faster adaptation. Mathematical modelling suggests that rapid adaptation of tetraploids is driven by higher rates of beneficial mutations with

stronger fitness effects, which is supported by whole-genome sequencing and phenotypic analyses of evolved clones. Chromosome aneuploidy, concerted chromosome loss, and point mutations all provide large fitness gains. We identify several mutations whose beneficial effects are manifest specifically in the tetraploid strains. Together, these results provide direct quantitative evidence that in some environments polyploidy can accelerate evolutionary adaptation.

To determine how polyploidy affects the rate of adaptation, we performed hundreds of independent passaging experiments in a poor carbon-source medium (raffinose, Fig. 1a), comparing isogenic haploid (1N), diploid (2N), and tetraploid (4N) strains (Extended Data Fig. 1 and Extended Data Table 1). The evolution experiments were performed as competitions between equal numbers of cells of the same ploidy expressing cyan fluorescent protein (CFP) or yellow fluorescent protein (YFP)^{17,18}, where the acquisition and spread of beneficial mutations is visualized by divergence from a 50:50 ratio of CFP- and YFP-expressing cells (Fig. 1b). The rate of adaptation was determined by measuring the change in

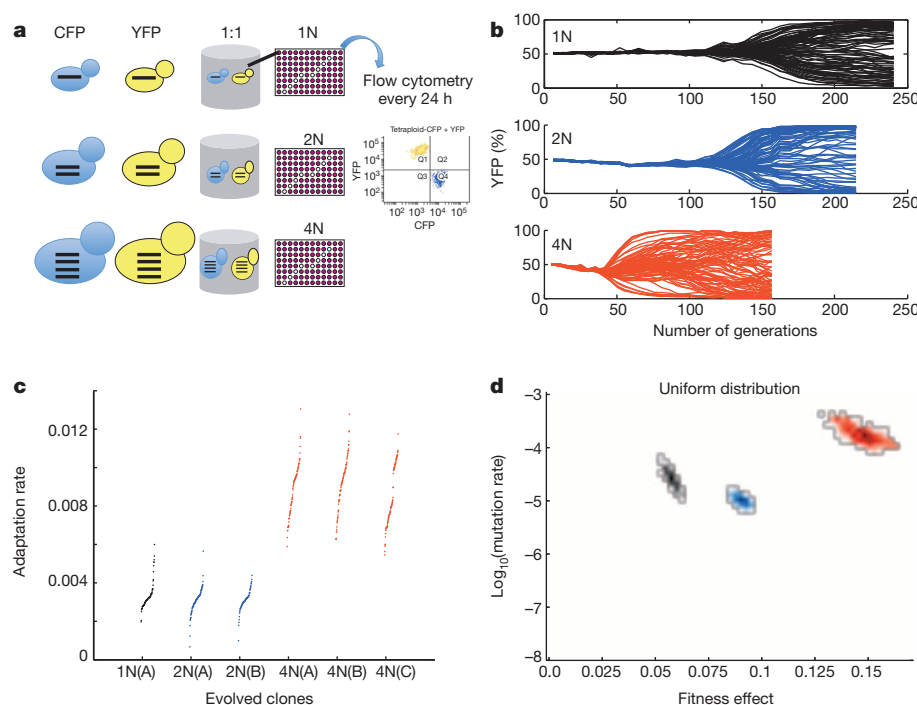


Figure 1 | Rapid spread of beneficial mutations in tetraploid yeast. **a**, Schematic diagram of the evolution experiment. **b**, Flow cytometry analysis of isogenic haploid (black), diploid (blue), and tetraploid (red) populations during adaptation to raffinose medium. Each line is the percentage of YFP cells in an independent population of YFP and CFP cells. Here and below, data from haploids is black, from diploids is blue, and from tetraploids is red. **c**, The adaptation rate of the evolved clones relative to the diploid ancestor after 250 generations. Data points are the average rate of adaptation (change in fitness between generation 250 and generation zero, divided by 250 generations) of two replicate fitness measurements for the evolved clones. Clones from replicate evolution experiments (A, B, and C) are indicated. The tetraploids acquired significantly more fitness in the same number of generations compared with the haploids and diploids (*t*-test, $P < 1 \times 10^{-10}$). **d**, Estimates from the branching evolution model of the best-fit value of the selection coefficient and beneficial mutation rate of each ploidy experiment, and their error range, determined using a uniform distribution of acquired mutations (other distributions are analysed in Extended Data Fig. 2c, d, and the equivalence principle model is analysed in Extended Data Fig. 2e). Error ranges were obtained by parametric bootstrap of 1,000 independent realizations (Methods).

¹Department of Pediatric Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, Massachusetts 02215, USA. ²Department of Cell Biology, Harvard Medical School, 25 Shattuck Street, Boston, Massachusetts 02215, USA. ³Howard Hughes Medical Institute, 4000 Jones Bridge Road, Chevy Chase, Maryland 20815, USA. ⁴Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, 450 Brookline Avenue Boston, Massachusetts 02215, USA. ⁵Department of Biostatistics, Harvard School of Public Health, 158 Longwood Avenue, Boston, Massachusetts 02215, USA. ⁶BioFrontiers Institute, University of Colorado at Boulder, 3415 Colorado Avenue, Boulder, Colorado 80303, USA. ⁷Department of Molecular, Cellular and Developmental Biology, University of Colorado at Boulder, 347 UCB, Boulder, Colorado 80309, USA. ⁸Broad Institute, 415 Main Street, Cambridge, Massachusetts 02142, USA. ⁹Department of Medicine, University of Colorado School of Medicine, 13001 East 17th Place, Aurora, Colorado 80045, USA. ¹⁰Department of Biostatistics and Informatics, Colorado School of Public Health, 13001 East 17th Place, Aurora, Colorado 80045, USA. ¹¹Molecular Oncology Program, University of Colorado Cancer Center, 13001 East 17th Place, Aurora, Colorado 80045, USA. ¹²Department of Systems Biology, Harvard Medical School, 200 Longwood Ave, Boston, Massachusetts 02115, USA. ¹³Department of Biology, Technion - Israel Institute of Technology, Haifa, 32000, Israel. ¹⁴Department of Pediatric Hematology/Oncology, Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, USA. †Present address: Department of Medical Microbiology and Immunology, Creighton University School of Medicine, 2500 California Plaza, Omaha, Nebraska 68178, USA.

fitness relative to the diploid ancestor over time (Methods). Over 250 generations, the tetraploids adapted at a rate that was significantly faster than haploids or diploids (Fig. 1c; t -test, $P < 1 \times 10^{-10}$, Methods). This faster rate of adaptation in tetraploids may be due to a higher rate of beneficial mutations, higher fitness effects of the acquired mutations, or both.

To gain insight into the rapid adaptation of tetraploids, we applied two complementary mathematical modelling approaches (see Methods). First, we use a model based on a branching evolutionary process¹⁹, designed to mimic closely the divergence experiments. At each time-step, a cell is chosen at random to die or to divide, with a probability corresponding to its fitness. Mutations arise with rate μ . If a mutation occurs, the fitness of the daughter cell may change and the fitness increase is then chosen from a fitness distribution. Second, we use the 'equivalence principle' model¹⁷, which focuses on beneficial mutations that establish in the population, and estimates that these mutations confer a single effective fitness advantage. Proliferation of clonal subpopulations under this model is deterministic. These simplifications are relevant to examples of high clonal interference, and are therefore only appropriate when the population size is large or when the beneficial mutation rate is high²⁰. In both models, we assume no epistasis; the fitness change is independent of whether the cell already had one or more mutations. Furthermore, there is no restriction on the number of cells that acquire beneficial mutations, thus allowing clonal interference to occur^{18,21}.

Both modelling approaches led to the same general conclusion: the rapid adaptation of tetraploids results from both more frequent beneficial mutations and stronger fitness effects (Extended Data Fig. 2 and Methods). For the branching evolution model, these conclusions are independent of the assumed distribution of beneficial mutations, although there are differences in the magnitude of the best-fit values that are expected from the shape of the chosen distribution (Fig. 1d and Extended Data Fig. 2). Moreover, the conclusions are insensitive to the inclusion of deleterious mutations in the model (Extended Data Fig. 3 and Methods).

To evaluate these conclusions experimentally, we performed whole-genome sequencing (WGS) to compare the frequency of mutations in the 1N and 4N ancestors with 74 evolved clones. In total, we identified 240 *de novo* sequence variants (single nucleotide polymorphisms (SNPs) and small insertions/deletions): 45 from the 1N-, 69 from the 2N-, and 126 from the 4N-evolved clones, an average of 2.05, 2.87, and 4.5 variants respectively per cell type (Supplementary Table 1). We observed significantly more variants per 4N clone than per 1N- and 2N-evolved clones (Fig. 2a; t -test, $P < 1 \times 10^{-4}$ and $P = 0.0040$, respectively). Note that these results are not a direct measurement of the mutation rate or beneficial mutation rate (μ), but rather the total number of mutations acquired during the experiment (see Supplementary Discussion).

Sequence variants frequently occurred in genes encoding proteins in the *Snf3/Rgt2* glucose-signalling pathway (*SNF3*, *RGT2*, *MTH1*, *RGT1*), as expected from previous yeast evolution experiments under carbon-source limitation^{22–24}. Several independent mutations in these genes resulted in either identical base-pair changes or altered the same amino acid (Supplementary Table 1). Non-synonymous *SNF3* mutations were identified in all ploidy types, whereas loss-of-function mutations in *MTH1* were observed most frequently in the 1N-evolved clones.

In addition to WGS, we used a combination of flow cytometry, microarray comparative genome hybridization (aCGH), and quantitative PCR (qPCR) to measure the frequency of DNA copy number variations (CNVs) in the evolved clones. The only CNVs that arose in all three ploidy types was amplification of two adjacent genes encoding the high-affinity hexose transporters, *HXT6* and *HXT7*, a frequently identified beneficial mutation in low glucose environments^{18,22,23}. The *HXT6/7* amplification was significantly more common in 2N- and 4N-evolved clones than in 1N clones (t -test, $P = 0.005$ and $P = 1 \times 10^{-4}$, respectively, Methods), which may be due to negative epistasis between *HXT6/7* amplification and mutations in 1N cells, such as those in *MTH1* (ref. 24).

Additional CNVs, including recurrent chromosome aneuploidy, were detected only in the 4N-evolved clones. With the exception of a small

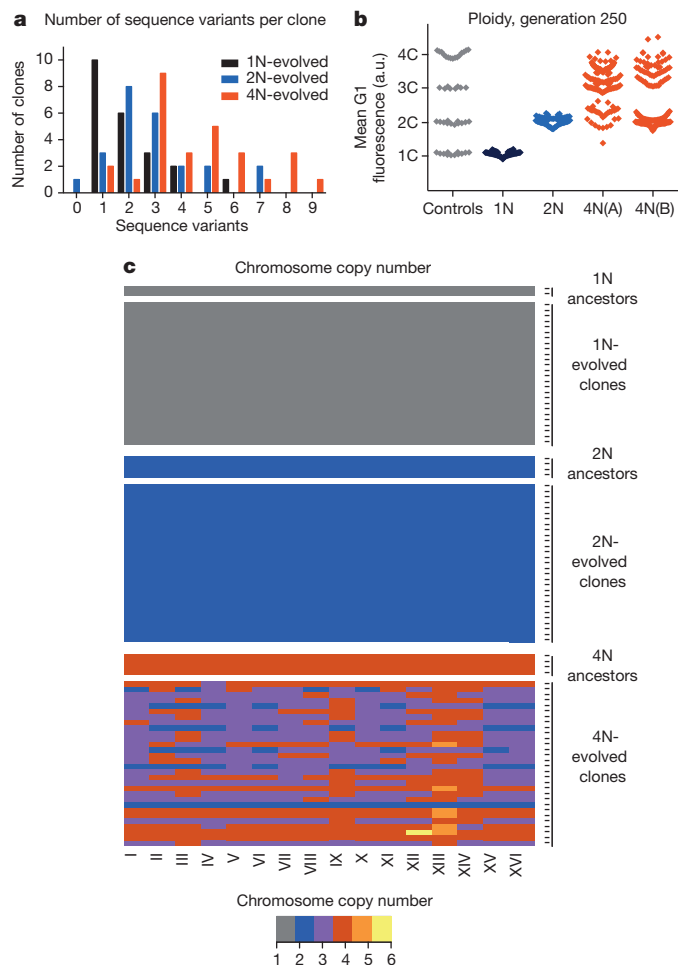


Figure 2 | Tetraploid clones acquire frequent sequence variants, large-scale ploidy shifts and recurrent whole chromosome aneuploidy during adaptation. **a**, The number of sequence variants per clone was determined with WGS of 74 evolved clones (22 haploid, 24 diploid, and 28 tetraploid clones; Supplementary Table 1). The difference between tetraploids and haploids or diploids was significant (t -test, $P < 1 \times 10^{-4}$ and $P = 0.004$, respectively). **b**, DNA content (1C–4C) of evolved clones at generation 250, measured as the mean G1 propidium iodide fluorescence for each evolved clone ($n = 88$ haploid and diploid evolved clones, $n = 176$ tetraploid evolved clones) in arbitrary units (a.u.). For reference, the DNA content of ancestral, control strains (1N, 2N, 3N, and 4N) is shown in grey. **c**, Heat map of chromosome copy number data obtained from aCGH and WGS for the ancestral and evolved 1N, 2N, and 4N clones at generation 250; colour key at bottom. See Extended Data Figs 4–6, and Supplementary Table 2 for all individual clones.

segmental amplification in one 2N-evolved clone, there were no CNVs or aneuploidy in the ancestral strains or the 1N- and 2N-evolved clones (Fig. 2b, c and Extended Data Figs 4 and 5). By contrast, many 4N clones underwent large shifts in ploidy (Fig. 2b) and all but two of the 4N-evolved clones were aneuploid at generation 250 ($n = 30$; Fig. 2c, Extended Data Fig. 6 and Supplementary Table 2). These alterations included large segmental aneuploidies with breakpoints at loci of transposable elements (Extended Data Fig. 7a). Pairwise patterns of chromosome copy number alterations were observed, indicating that there is a strong copy number relationship between certain pairs of chromosomes (Extended Data Fig. 7b, c). Notably, increased copy number of chromosome XIII was significantly more common than all other aneuploidies (Extended Data Fig. 7d; Cochran–Armitage test, $P < 1 \times 10^{-7}$). These chromosome-level alterations were present early, at the time of CFP/YFP marker divergence in the 4N populations (\sim generation 45; Extended Data Fig. 8). Therefore, 4N-evolved clones had a higher frequency and greater diversity of mutations, supporting the inference from our mathematical

model that 4N-evolved clones have a relatively higher beneficial mutation rate.

Next, we determined the effects of specific mutations on the fitness of the ancestral cells of differing ploidy. We first determined whether chromosome XIII gain contributed directly to the rapid adaptation of 4N cells. Isogenic 2N and 4N strains, with and without an extra copy of chromosome XIII, were generated (Methods and Extended Data Fig. 9). The increased copy number of chromosome XIII provided a significant fitness increase to 4N strains specifically in raffinose medium relative to the 2N ancestor (Fig. 3a; *t*-test, $P < 1 \times 10^{-4}$), not in glucose (Fig. 3b). This was not a general effect of aneuploidy because the gain of a different chromosome, chromosome XII, had the opposite effect on fitness (Fig. 3a). In striking contrast to 4N cells, chromosome XIII trisomy was not beneficial to 2N strains in raffinose medium and decreased fitness of 2N cells in glucose. Although increased fitness due to whole and segmental chromosome gain is known to occur during adaptation^{13,14,23}, to our knowledge this is the first observation of a ploidy-specific fitness advantage for an aneuploid chromosome. Thus, aneuploidy, acquired through high rates of mitotic errors, is one way that 4N cells can acquire more beneficial mutations with higher fitness effects.

We also characterized how ploidy impacts the fitness effect of recurrently isolated mutations in *SNF3*, a gene encoding a plasma-membrane glucose sensor²⁵. We identified SNPs that changed the codon for the same amino acid in the ninth transmembrane domain of Snf3p (G439E, G439V, G439R; Supplementary Table 1), and increased *HXT* expression in raffinose (Methods)²⁵. By analysing the fitness of isogenic *SNF3*–*G439E* strains differing only by ploidy, we found that *SNF3*–*G439E* had a dominant, raffinose-specific, beneficial effect that was relatively stronger in the 4N strain (Fig. 3c, d; *t*-test, $P < 1 \times 10^{-4}$).

Ploidy-specific effect size of mutations could be an intrinsic property of polyploidy, as was recently suggested in plants¹⁶, or it could be related

to the fitness of the 4N ancestor relative to the 1N and 2N ancestors^{26,27}. To address the impact of initial fitness generally, we isolated 48 clones from the 4N evolution experiments at generation 250 (4N₂₅₀) with fitness values equal to the 2N ancestor (competitive fitness difference < 0.05), and determined the speed of their next adaptive step. We compared the fitness acquired by the selected 4N₂₅₀ clones after an additional 250 generations with that of 2N clones evolved for 250 generations (2N₂₅₀, $n = 160$). Despite comparable starting fitness, the 4N-derived clones still underwent more rapid adaptation and achieved significantly higher fitness. This occurred irrespective of large-scale shifts in ploidy: 29% of the 4N₅₀₀ clones maintained a ploidy of 3N–4N and acquired higher fitness than the 2N₂₅₀ clones (Fig. 4, Kolmogorov–Smirnov test, $P < 1 \times 10^{-6}$); 71% of the 4N₅₀₀ clones underwent chromosome loss to become near-diploid and acquired even higher fitness relative to 2N₂₅₀ clones (Fig. 4, Kolmogorov–Smirnov test, $P < 1 \times 10^{-8}$). Thus, the rapid adaptation of tetraploid cells was at least partly independent of their initial fitness.

Here, we measured the acquisition and spread of beneficial mutations in isogenic yeast populations that differed only by ploidy. Mathematical modelling enabled us to infer parameters driving the evolutionary dynamics of these strains and indicated that in a poor carbon-source environment, polyploidy increases the rate and fitness effects of the acquired mutations. Polyploidy increased the genetic diversity of the population. We identified examples of mutations that are selectively beneficial in polyploid strains, including whole chromosome aneuploidy. Because aneuploidy itself is mutagenic²⁸, the high rates of aneuploidy induced by whole-genome duplication may further increase the rate at which beneficial mutations are acquired. If these mutations are beneficial at lower ploidy states, then the long-term benefit of polyploidy will be preserved, even if polyploidy is transient during adaptation. Indeed, 4N-evolved clones that became near-diploid had higher fitness than the 2N-evolved clones. Moreover, although we only studied one environmental condition, polyploidy buffers the effects of partly recessive deleterious mutations^{12,29}, which in principle can then accumulate², providing a reservoir of mutations that might be adaptive in a new environment. Interestingly, the evolved tetraploid karyotypes closely resemble the polyploid and aneuploid karyotypes of fermentation, industrial, baking, natural desert isolates³⁰, and antifungal drug-resistant yeasts¹⁴, consistent with a role for polyploidization events during adaptation to

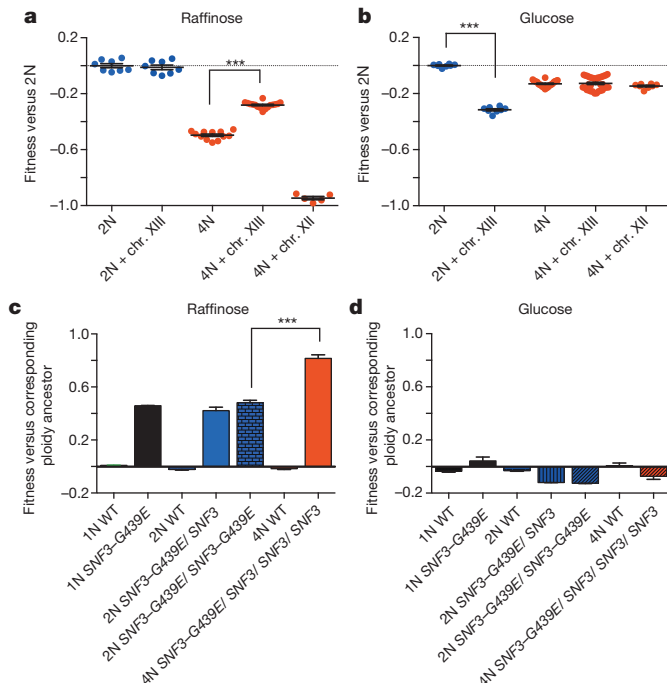


Figure 3 | Ploidy-specific fitness effects for certain beneficial mutations. Gain of chromosome (chr.) XIII is beneficial to tetraploid cells grown in raffinose medium but not for diploids. Shown is the fitness of isogenic wild-type 2N and 4N strains, with or without chromosome XIII gain, relative to the 2N ancestor in raffinose (a) or glucose (b) medium. Error bars, mean \pm s.e.m. of four individual clones and two or three technical replicates. c, Competitive fitness of engineered isogenic strains of the indicated ploidy and genotype, relative to the corresponding ploidy ancestor, in raffinose and d, glucose medium. Error bars, mean \pm s.e.m. of three independent *SNF3*–*G439E* transformants of each ploidy type, *t*-test *** $P < 1 \times 10^{-4}$.

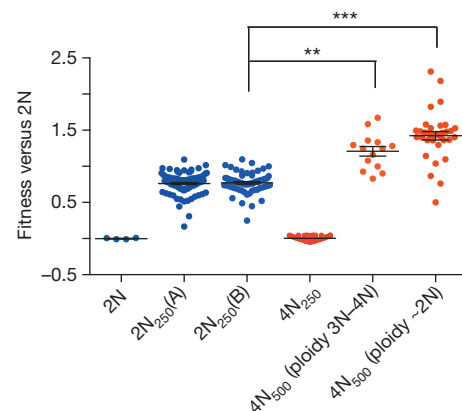


Figure 4 | Rapid adaptation of tetraploids normalized for initial fitness. Fitness of 2N and 4N clones relative to the 2N ancestor. Evolved tetraploids (4N₂₅₀) with fitness equivalent to the diploid ancestors were identified and passed for another 250 generations to generate 4N₅₀₀ clones ($n = 48$). The fitness of these 4N₅₀₀ clones was then compared with the fitness of evolved diploids after 250 generations (2N₂₅₀, $n = 160$, replicate experiments A and B). 4N₅₀₀ clones reached a higher fitness than 2N₂₅₀ clones, irrespective of whether the 4N₅₀₀ clones maintained a 3N–4N DNA content ($n = 14$, Kolmogorov–Smirnov test, ** $P < 1 \times 10^{-6}$) or underwent large-scale chromosome loss to a near diploid chromosome content ($n = 34$, Kolmogorov–Smirnov test, *** $P < 1 \times 10^{-8}$). Error bars, mean \pm s.e.m.

these stressful environments. Thus, the genetic plasticity of polyploid cells, together with ploidy-specific beneficial effects, can facilitate rapid adaptation.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 10 January; accepted 29 December 2014.

Published online 2 March 2015.

- Ohno, S., Wolf, U. & Atkin, N. B. Evolution from fish to mammals by gene duplication. *Hereditas* **59**, 169–187 (1968).
- Otto, S. P. & Whitton, J. Polyploid incidence and evolution. *Annu. Rev. Genet.* **34**, 401–437 (2000).
- Semon, M. & Wolfe, K. H. Consequences of genome duplication. *Curr. Opin. Genet. Dev.* **17**, 505–512 (2007).
- Hufton, A. L. & Panopoulou, G. Polyploidy and genome restructuring: a variety of outcomes. *Curr. Opin. Genet. Dev.* **19**, 600–606 (2009).
- Paquin, C. & Adams, J. Frequency of fixation of adaptive mutations is higher in evolving diploid than haploid yeast populations. *Nature* **302**, 495–500 (1983).
- Anderson, J. B., Sirjusingh, C. & Ricker, N. Haploidy, diploidy and evolution of antifungal drug resistance in *Saccharomyces cerevisiae*. *Genetics* **168**, 1915–1923 (2004).
- Zorgo, E. *et al.* Ancient evolutionary trade-offs between yeast ploidy states. *PLoS Genet.* **9**, e1003388 (2013).
- Mayer, V. W. & Aguilera, A. High levels of chromosome instability in polyploids of *Saccharomyces cerevisiae*. *Mutat. Res.* **231**, 177–186 (1990).
- Bennett, R. J. & Johnson, A. D. Completion of a parasexual cycle in *Candida albicans* by induced chromosome loss in tetraploid strains. *EMBO J.* **22**, 2505–2515 (2003).
- Gerstein, A. C., Chun, H. J., Grant, A. & Otto, S. P. Genomic convergence toward diploidy in *Saccharomyces cerevisiae*. *PLoS Genet.* **2**, e145 (2006).
- Fujiwara, T. *et al.* Cytokinesis failure generating tetraploids promotes tumorigenesis in p53-null cells. *Nature* **437**, 1043–1047 (2005).
- Storchova, Z. *et al.* Genome-wide genetic analysis of polyploidy in yeast. *Nature* **443**, 541–547 (2006).
- Rancati, G. *et al.* Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell* **135**, 879–893 (2008).
- Selmecki, A., Forche, A. & Berman, J. Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* **313**, 367–370 (2006).
- Zack, T. I. *et al.* Pan-cancer patterns of somatic copy number alteration. *Nature Genet.* **45**, 1134–1140 (2013).
- Chao, D. Y. *et al.* Polyploids exhibit higher potassium uptake and salinity tolerance in *Arabidopsis*. *Science* **341**, 658–659 (2013).
- Hegreness, M., Shores, N., Hartl, D. & Kishony, R. An equivalence principle for the incorporation of favorable mutations in asexual populations. *Science* **311**, 1615–1617 (2006).
- Kao, K. C. & Sherlock, G. Molecular characterization of clonal interference during adaptive evolution in asexual populations of *Saccharomyces cerevisiae*. *Nature Genet.* **40**, 1499–1504 (2008).
- Haccou, P., Jagers, P., Vatutin, V. A. & International Institute for Applied Systems Analysis. *Branching Processes: Variation, Growth, and Extinction of Populations* 82–106 (Cambridge Univ. Press, 2005).
- Barrett, R. D., M'Gonigle, L. K. & Otto, S. P. The distribution of beneficial mutant effects under strong selection. *Genetics* **174**, 2071–2079 (2006).
- Lang, G. I. *et al.* Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500**, 571–574 (2013).
- Brown, C. J., Todd, K. M. & Rosenzweig, R. F. Multiple duplications of yeast hexose transport genes in response to selection in a glucose-limited environment. *Mol. Biol. Evol.* **15**, 931–942 (1998).
- Gresham, D. *et al.* The repertoire and dynamics of evolutionary adaptations to controlled nutrient-limited environments in yeast. *PLoS Genet.* **4**, e1000303 (2008).
- Kvitek, D. J. & Sherlock, G. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet.* **7**, e1002056 (2011).
- Ozcan, S., Dover, J., Rosenwald, A. G., Wolf, S. & Johnston, M. Two glucose transporters in *Saccharomyces cerevisiae* are glucose sensors that generate a signal for induction of gene expression. *Proc. Natl Acad. Sci. USA* **93**, 12428–12432 (1996).
- Barrick, J. E., Kauth, M. R., Streltsov, C. C. & Lenski, R. E. *Escherichia coli* rpoB mutants have increased evolvability in proportion to their fitness defects. *Mol. Biol. Evol.* **27**, 1338–1347 (2010).
- Kryazhimskiy, S., Rice, D. P., Jerison, E. R. & Desai, M. M. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* **344**, 1519–1522 (2014).
- Sheltzer, J. M. *et al.* Aneuploidy drives genomic instability in yeast. *Science* **333**, 1026–1030 (2011).
- Dewhurst, S. M. *et al.* Tolerance of whole-genome doubling propagates chromosomal instability and accelerates cancer genome evolution. *Cancer Disc.* **4**, 175–185 (2014).
- Ezov, T. K. *et al.* Molecular-genetic biodiversity in a natural population of the yeast *Saccharomyces cerevisiae* from “Evolution Canyon”: microsatellite polymorphism, ploidy and controversial sexual status. *Genetics* **174**, 1455–1468 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements This work was supported by the Howard Hughes Medical Institute, the National Institutes of Health (R37 GM61345), the G. Harold & Leila Y. Mathers Charitable Foundation, the Dana-Farber Cancer Institute Physical Sciences-Oncology Center (U54CA143798), the Boettcher Foundation's Webb-Waring Biomedical Research Program, the National Science Foundation (NSF 1350915), the National Institutes of Health (R01 GM081617), and an American Cancer Society Postdoctoral Fellowship.

Author Contributions A.M.S., M.G., N.S., R.K., and D.P. contributed to the overall study design. A.M.S. and M.G. performed the experiments. Y.E.M. implemented the mathematical modelling with contributions from N.S., R.K., and F.M. A.M.S., P.A.R., and A.L.S. generated WGS libraries; P.A.R. developed the sequencing pipeline and analysed the WGS data with help from A.M.S. and A.L.S. under the supervision of R.D.D. Data were analysed by A.M.S., Y.E.M., P.A.R., M.G., N.S., A.L.S., S.D., F.M., and D.P. The manuscript was written primarily by A.M.S., Y.E.M., and D.P. with contributions from the other authors.

Author Information All aCGH data have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus database under accession number GSE51017 and all WGS data have been deposited in the National Center for Biotechnology Information Sequence Read Archive database under accession number SRP047435. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.M.S. (AnnaSelmecki@creighton.edu) or D.P. (david_pellman@dfci.harvard.edu).

METHODS

Batch culture evolution experiment. All *Saccharomyces cerevisiae* strains used in this study were in the S288c background (detailed information on strain construction is provided in the section entitled 'Yeast strain construction'). Briefly, the isogenic ploidy series was generated in a *matA ste4A* background to eliminate mating and meiosis during the course of the experiment. Either a *pGAL-CFP* or a *pGAL-YFP* construct was integrated at the *TRP1* locus near the chromosome IV centromere in a haploid strain (PY5998 and PY5999, respectively). These haploid strains were used to generate isogenic diploids, from which isogenic tetraploids were then derived (Extended Data Fig. 1). This procedure ensured that all copies of chromosome IV had the capacity to express the inducible fluorescent marker even if the strains became aneuploid. Mating-competent haploids were generated from the *matA ste4A* ancestor, PY5998, by transformation with either plasmid PB2647 (*CEN-LEU2-STE4*) or PB2648 (*CEN-URA3-STE4-Matz*). Zygotes from mating-competent haploids were isolated by micromanipulation to obtain diploid CFP ancestors (PY6008 and PY6022). Similarly, zygotes from mating-competent diploids were isolated by micro-manipulation to obtain tetraploid CFP ancestors (PY6031 and PY6032). The same mating scheme was performed for the YFP lineage starting with PY5999 to generate diploid YFP (PY6006 and PY6014) and tetraploid YFP (PY6040 and PY6045) ancestors.

The ancestor strains were grown to saturation from the -80°C stock, in Synthetic Complete + 2% glucose. The cell density of each ancestor was determined using a haemocytometer and an automated cell counter (Vi-Cell-XR from Beckman Coulter). An equal number of YFP and CFP cells of the same ploidy were diluted into fresh Synthetic Complete + 2% raffinose medium, and combined into a single tube for an initial concentration of 1×10^5 cells per millilitre. The 50:50 YFP:CFP culture was distributed equally into the wells of a 96-deep-well plate (1 ml per well, U-bottom block plate from Qiagen). Seven or eight wells were not inoculated, to detect cross-well contamination during the experiment. The plates were covered with 'breathe-EASIER' tape (Electron Microscopy Science) and incubated at 30°C on a 96-well plate shaker (Union Scientific). Two plates of haploid and three plates of diploid and tetraploid cells were analysed, representing 173 parallel haploid evolutions, 264 parallel diploid evolutions, and 265 parallel tetraploid evolutions.

At 24 h intervals, the cells were resuspended (by pipetting) and diluted into fresh Synthetic Complete + 2% raffinose medium. The dilution factor was determined for each ploidy type based on the initial strain fitness to maintain an equivalent population size, as reported previously¹⁷. The number of cells transferred each day was calculated by counting the number of cells in ten replicate wells of each ploidy before and after dilution with an automated cell counter (Vi-Cell-XR from Beckman Coulter), and averaged across 3 consecutive days. The dilution factor for the haploid, diploid, and tetraploid experiments was 1/100, 1/50, and 1/33, respectively. This corresponds to 6.64, 5.64, and 5.04 generations per day¹⁷. The tetraploid evolution experiment from generation 250 to 500 (Fig. 4) was performed with the same dilution factor as the diploid experiments (1/50).

The number of CFP and YFP cells in each population was measured at the same time each day. First, expression of the fluorescent proteins was induced by transferring 10 μl of the overnight culture into 200 μl Synthetic Complete + 2% galactose medium for 4 h at 30°C . The number of CFP- or YFP-expressing cells was determined using a BD LSRII flow cytometer high-throughput plate reader (10,000 cells were analysed from each well). Pacific Blue and FITC filters were used to detect CFP and YFP, respectively. All experiments were passaged for 250 generations, but daily acquisition of CFP:YFP ratios was not always continued to the 250th generation.

To ensure that the flow cytometer measurement and the galactose induction of CFP and YFP was an accurate reflection of the size of these populations, the ratio of CFP:YFP cells was determined by both flow cytometry and microscopy, and the ratio was determined both before and after galactose induction. To do this, we combined overnight cultures of the 1N, 2N, and 4N ancestor CFP and YFP strains at three different ratios (nine populations in total) and analysed the ratios in two ways. First, for an aliquot of the mixture, we induced the expression of the fluorescent proteins with 2% galactose for 4 h and analysed 10,000 cells using flow cytometry. In parallel, we also added 2% galactose for 4 h and then counted about 300 cells by fluorescence microscopy. Finally, to ensure that the induction with 2% galactose did not alter the CFP:YFP ratio, a portion of the population was used to determine the number of CFP and YFP cells in the population before adding galactose to the medium. To do this, cells from each population were struck for single colonies on YPD plates for 2 days. Ninety-six colonies were chosen randomly from each plate and added to a single well of a 96-well plate containing Synthetic Complete + 2% galactose. The fluorescence of each colony was determined by flow cytometry, and the percentage YFP of the initial population was determined. There was a strong correlation between the percentage YFP-expressing cells obtained from all three measurements (Extended Data Fig. 3a): including the flow cytometer and fluorescence microscopy (Pearson correlation coefficient = 0.979), and both before and after galactose induction (Pearson correlation coefficient = 0.985).

Finally, frozen stocks of the evolution experiments were made at 3- to 4-day intervals throughout the experiment. At the end of each experiment, single colony clones were isolated and used for competitive fitness assays, flow cytometry analysis of ploidy, and preparation of DNA for aCGH.

We isolated 48 clones from the 4N evolution experiments at generation 250 ($4N_{250}$) with fitness values equal to the 2N ancestor (competitive fitness difference < 0.05), and determined the rate of adaptation after an additional 250 generations (Fig. 4). Each $4N_{250}$ clone was grown to saturation from the -80°C stock in Synthetic Complete + 2% raffinose medium. Cell counts were performed as above, and each population was diluted to an initial concentration of 1×10^5 cells per millilitre. At 24 h intervals, the cells were resuspended (by pipetting) and diluted 1/50 into fresh Synthetic Complete + 2% raffinose medium (the same dilution factor as the diploid experiments). These evolution experiments were not performed as CFP:YFP competitions, so daily flow cytometry was not necessary. After 250 generations, single colony clones ($4N_{500}$) were isolated on Synthetic Complete + 2% raffinose plates. Each $4N_{500}$ clone was cultured overnight in 1 ml Synthetic Complete + 2% raffinose medium and aliquots of this culture were immediately used for competitive fitness assays, flow cytometry analysis of ploidy, preparation of DNA for aCGH, and frozen stocks.

Measuring the variation in the flow cytometer measurements. We determined the amount of noise in our flow cytometer measurements by calculating the mean and standard deviation of the percentage YFP obtained from 48 independent populations at six different ratios of CFP:YFP, for each ploidy type. The 1N, 2N, and 4N ancestor strains were cultured separately overnight in 2% raffinose medium and transferred to 2% galactose for 4 h to induce expression of CFP and YFP. Next, CFP and YFP cells of the same ploidy were combined at ratios of 100:0, 85:15, 75:25, 50:50, 25:75, and 0:100 to reach the same final volume (200 μl). Ten thousand cells from each population were analysed by the LSRII (BD) flow cytometer using the same parameters (for example gating and flow rate) that we used for the evolution experiments, the total numbers of CFP and YFP cells were obtained, and the percentage YFP was calculated ((number of YFP cells)/(number of CFP + number of YFP cells)) $\times 100$. The standard deviations are presented in Extended Data Fig. 3b, and indicate that there was little well-to-well variability for the same CFP:YFP ratio (across 48 wells), and that this variability changed only slightly across different CFP:YFP ratios of all three ploidy types. Importantly, it was never greater than 0.66% of the measurement. This small variability had a minimal effect on our analysis because the fitting procedure we used to measure the deviation from an equal percentage of CFP and YFP cells used bins of 5% deviation to combine the number of wells that had deviations between 0 and 5%, 5 and 10%, and so on, for each experiment.

Statistical analysis of the experimental data. No statistical methods were used to predetermine sample size.

Adaptation rate. The CFP versus YFP evolution experiments were designed to analyse the dynamics of the adaptation process at the population level within and across wells. The adaptation rate was determined with additional competition experiments that were designed to measure the change in fitness from generation zero to generation 250 for cells of each ploidy. To this end, we isolated single colony clones from each evolved well of different ploidy types at generation 250 and measured their competitive fitness relative to the 2N ancestor (see detailed methods below). Fitness was defined as the slope of the log of the ratio of the evolved clone to the reference strain (2N ancestor) over time; more precisely, we performed a linear least-squares fit of $\log(N_{t1}/N_{t0})$ over multiple dilution cycles (where N_{t1} is the number of cells from the evolved clone and N_{t0} is the number of ancestor cells). The fitness relative to the ancestor is defined as $s = d/dt [\log_2(N_{t1}/N_{t0})]$, where t is measured in days¹⁷. The rate of adaptation was the relative fitness at generation 250 minus the relative fitness of generation zero, divided by 250 generations. The rate of adaptation for each ploidy is shown in Fig. 1c. We found that the tetraploid populations had a significantly larger rate of adaptation (0.009 [0.0062, 0.011]) than haploids (0.0031 [0.0022, 0.0052]) or diploids (0.0031 [0.0018, 0.0041]) during the 250 generations in raffinose medium (values indicate the median rate followed by the 95% confidence intervals in square brackets, Fig. 1c, t -test, $P < 1 \times 10^{-10}$).

Mathematical modelling of population dynamics. Branching evolution model. We formulated a mathematical model of the population dynamics of cells that was then used to infer evolutionary parameters using the experimental data. This branching evolution model was designed to mimic closely the divergence experiments containing two equally fit populations (CFP or YFP), each initially consisting of 50,000 cells. The model is based on a stochastic birth and death process called a branching process³¹. In this process, at each time-step, a cell is chosen to die at random, or to divide with a probability corresponding to its fitness. During each cell division, a mutation arises with mutation rate μ . If no mutation occurs, the fitness of the daughter cell is equal to the fitness of the mother cell. If a mutation does occur, the fitness may change; the additive fitness of the daughter cell is then chosen from a fitness distribution. The fitness change is independent of whether the ancestors of this clone had already obtained one or more mutations. Furthermore, there

is no restriction on the number of cells that acquire beneficial mutations, thus allowing clonal interference to occur^{18,21}.

We compared results assuming either a uniform, exponential, or delta distribution of fitness values (Extended Data Fig. 2a–d). For the initial formulation of the model, we considered half of the newly arising mutations to be beneficial and half deleterious, and their fitness effects were considered to be additive to the fitness value of the mother cell. Because complete simulations of the branching process would be prohibitively slow, we approximated this branching process with a Wright–Fisher process¹⁹ with non-overlapping generations. This process was implemented as a Monte Carlo simulation in C++ and the code is provided as a Supplementary Software file.

In the simulation, each competition experiment was initiated with 1×10^5 cells (5×10^4 of each CFP and YFP cell type). At every generation, each cell reproduces and gives birth to a random number of surviving offspring distributed according to a Poisson distribution. The initial population of each ploidy had a different growth rate per day: the haploid population increased approximately 100-fold, the diploids about 50-fold, and the tetraploids about 30-fold in a 24 h time interval (described previously in the section entitled ‘Batch culture evolution experiment’). Assuming a population doubling every generation, the average number of generations per day was 6.64, 5.65, and 5.04 for haploids, diploids, and tetraploids, respectively. In the Wright–Fisher model, the number of generations is discrete; thus we rounded these numbers to the closest integers (7, 6, and 5, respectively). The initial fitness (f) of each ploidy was chosen to satisfy the growth rate of that ploidy with this number of generations: that is, $(f_{1N})^7 = 100$, $(f_{2N})^6 = 50$, $(f_{4N})^5 = 30$. The fitness of a cell is the average number of its surviving offspring in the next generation.

During each cell division, a new mutation might arise. In haploid cells, the impact of a mutation is given by its additive fitness value s drawn from the fitness distribution. In diploid and tetraploid cells, however, a mutation might have a degree of dominance, be recessive or have a different fitness effect than it would have in a haploid cell. Following ref. 2, we assumed that the effect of a given mutation in a diploid or tetraploid cell compared with that in a haploid cell is scaled by a certain factor h . Therefore, the fitness effect of a given mutation in one allele will be $\tilde{s} = hs$.

In our model, we considered a beneficial mutation to arise independently at rate μ (per whole genome). Because our populations proliferated asexually, we could not differentiate between the two components of \tilde{s} , namely the haploid fitness effect s and the dominance coefficient h , and therefore we could only infer the combined value of \tilde{s} . Our simulations were done in exactly the same way for all ploidy types, and the selection coefficient of a new mutation was taken from a given distribution. However, the meaning of the added fitness is different for the haploids, where it is s , than for the diploids and tetraploids, where it is \tilde{s} . The assumption of independence of mutations is justified by the low point mutation rate, which is of the order of magnitude of 2×10^{-10} per base per generation in yeast³². The low per-base mutation rate means that the probability of independently obtaining a second identical mutation is vanishingly small (probability of μ^2 , or 4×10^{-20}), and therefore we did not consider such events in the model. Similarly, the probability of obtaining a given mutation that is then copied by a recombination-based mechanism such as gene conversion³³ is low ($(2 \times 10^{-10}) \times (4 \times 10^{-5})$), and was not included. These assumptions are validated by our WGS data demonstrating single copies of all point mutations.

Our identification of specific mutations that have a larger fitness effect in the tetraploid strains than in the haploid or diploid strains (Fig. 3) is consistent with the overall larger fitness effect of the tetraploids in our experiment. Note, however, that our parameters describe the entire distribution of mutations, rather than any specific mutation. Thus, it is also possible that some mutations could have a larger fitness effect in the haploids or diploids relative to the tetraploids. Note also that in our model clonal interference can occur, as there is no limitation on the number of independent mutations that can arise within a population. Thus multiple clones can emerge and compete with each other. As shown previously, clonal interference is an important aspect of microbial population dynamics and thus cannot be ignored^{17,18,21}. Additionally, multiple mutations can exist in the same cell.

As a sensitivity analysis, we also varied the ratio between advantageous and deleterious mutations, and found that even when the fraction of the deleterious mutation is very large, their inclusion has only a negligible effect on the rate and dynamics of adaptation. This is consistent with a large body of prior literature^{17,34–36}. To accomplish this, we fixed the rate of the advantageous mutations and generated data sets with different ratios of beneficial to deleterious mutations: 100:0, 90:10, 50:50, 10:90, and 1:99. For each ratio we generated 100 data sets (each data set with 264 single deviation experiments) using a beneficial mutation rate of 1.2×10^{-6} and a fitness effect of $s = 0.16$ (the best-fit values of the delta function for the diploid experiments, as a typical example, we used the same distribution as in ref. 17), and fitted them against the simulations that we used to fit the empirical experiments (see below for the fitting procedure). The means of the fitted values are presented in Extended Data Fig. 3c. We found no significant difference between the different

ratios of beneficial to deleterious mutations (t -test, $P > 0.2$ between all ratios tested). Therefore, in our model, we assumed that half of all non-neutral mutations were deleterious and half were beneficial, and that the deleterious and beneficial mutations had the same fitness effect distribution with the same parameters. Note that we also obtained similar results using the equivalence principle model¹⁷ that did not include deleterious mutations (see below).

The upper bound of cellular fitness. For biological plausibility, it was necessary to set an upper bound for acquired fitness. On the basis of the well-described growth rates achievable for *S. cerevisiae* in optimal conditions, we set this boundary at a doubling time of 1 h (2^{24} per day). Given that g is the number of generations a strain experiences in a day, the upper bound of the fitness f of any ploidy type was therefore set to satisfy the equation $f^g = 2^{24}$.

The initial growth rate of the tetraploid cells was lower than the growth rate of the haploids and diploids (see previous section entitled ‘Batch culture evolution experiment’). In the simulations we rounded the number of generations to satisfy the assumption of non-overlapping generations in the Wright–Fisher model: the tetraploid cells underwent about five generations per day, whereas the haploid cells underwent around seven, and diploid cells around six generations per day. In the simulations, as in the experiments, we diluted the populations every day by choosing at random 1% of the haploid, 2% of the diploid, and 3.3% of the tetraploid cells. This dilution was done by using a hypergeometric random generator for populations smaller than 100 million (<http://www.agner.org/random/>); for populations larger than 100 million this method is not applicable and we used a direct Bernoulli sampling of cells³⁷, forcing the total sampled cells to be 1%, 2%, or 3.3% for haploid, diploid, and tetraploid populations, respectively. The concentration of each of the cell types was recorded in the simulation output at the end of every day, after dilution.

In the experiments with the diploid and tetraploid cells, we observed an initial small bias against the YFP-labelled population: in the diploid cells there was a decline in the YFP-labelled population of 0.5% per day, and in the tetraploid cells there was a 1% decline per day (Fig. 1b). We included these biases in the simulations by including this initial small difference in the fitness (that is, a reduced average number of offspring) of CFP- versus YFP-labelled cells^{17,26}.

Note that drift could not feasibly play a role in our experiment, as the timescale (in generations) for fixation of an allele due to drift is approximately equal to the effective population size³⁸. This time was of the order of 10^6 generations in our experiments¹⁷, much larger than the timeframe of our experiment (250 generations). Thus the simulations continued until one of the cell types overtook the whole population or until the end of the time that the CFP:YFP data were collected on the 30th day; the first of either event terminated a simulation. To increase the efficiency of the simulations, extinction of a certain colour (and thus fixation of the other) was defined not by a value of zero frequency, but as a frequency of less than or equal to 1%, as the probability was negligible that a sub-clone present at a frequency of 1% would overtake another sub-clone present at a frequency of 99% in the timeframe of our experiment²¹. Furthermore, while the variability between flow cytometer measurements was never greater than 1% of the measurement (see above, and ref. 26), our ability to detect changes in CFP or YFP populations below a frequency of 1% was a limitation of the flow cytometer.

Fitting procedure. The empirical data for each ploidy type were combined to represent the average deviation from an equal percentage of YFP and CFP cells. We then used the combined ploidy data to estimate the best-fit values of mutation rate (μ) and selection coefficient (s), using least squares fitting.

To compare data from the experiment and the simulations, we investigated several different summary statistics and used the one that performed best. The summary statistics evaluated were as follows.

1. Mean deviation from an equal percentage of CFP and YFP cells: for every experiment corresponding to an individual well in a 96-well plate, we calculated the deviation from equal percentage every day. Then, for every data set generated (experiment and simulation) we calculated the average deviation per day. This procedure created a vector of the mean deviations per day for every data set.
2. Mean and standard deviation of the deviation from an equal percentage of CFP and YFP: as in 1, except that we also calculated the standard deviation of the deviation from equal percentage for every day. This procedure generated two vectors for every data set.
3. A distribution of the deviation from an equal percentage of CFP and YFP with ten bins: we calculated the absolute value of the deviation from equal percentage for each deviation experiment for each day. Then for every day, we binned the deviation values into ten bins, each with a size of 5%. That is, we counted the number of wells that had a deviation between 0–5%, 5–10%, ..., 45–50%. This procedure generated a matrix of ten bins \times 30 days for each data set.
4. A distribution of the deviation from equal percentage CFP and YFP with three bins: similar to (3), except we used only three bins, one for the non-deviated wells (defined as 0–0.1), one for fixated wells (defined as 0.4–0.5) and one for those in the middle (0.1–0.4).

The comparison between the experiments and the simulations was done by calculating the sum of squares (SOS) between the summary statistic (SS) of the experiment and the summary statistic of the simulations:

$$\text{SOS}_{(\mu,s)} = \sum_{\Omega} \left[\left(\text{SS}_{(\mu,s)}^{\text{Experiment}}(\Omega) - \text{SS}_{(\mu,s)}^{\text{Simulation}}(\Omega) \right)^2 \right]$$

where Ω spans the day and the number of values that each summary statistic has each day. The best-fit pair was the set of μ and s values with the smallest sum of squares. To determine the performance of the different summary statistic, we generated 1,000 data sets with the same parameters (1N, uniform distribution with $s = 0.05$, and $\mu = 8 \times 10^{-5}$), each containing 264 single deviation experiments, and inferred their values by scanning the parameter space for each of the 1,000 artificial data sets. The expected values from the simulations were then generated by 1,000 single deviation experiments for each set of parameters.

The distribution of inferred μ values is shown in Extended Data Fig. 3d. Whereas all of the summary statistics had the mutation rate used in the simulations as their mode, the observed means varied modestly. The ten-bin summary statistic had the narrowest range, and therefore this was used for further analysis. A similar summary statistic was also used previously³⁹.

The scanned range of mutation rates and fitness effects. The best-fit value was found by scanning a range of mutation rates and fitness effects²⁶; to this end, we scanned the parameter space of the fitness effect in linear steps from 0.005 to 0.35, with increments of 0.005. The mutation rates were scanned in logarithmic steps from $\log_{10}(\mu) = -8$ to $\log_{10}(\mu) = -4$, with increments of 0.1. Thus, per ploidy per fitness distribution, we scanned 2,460 parameter regimes, analysed 1,000 divergence simulations per μ and s combination, and scanned three ploidy types and three different fitness distributions in total.

For the exponential fitness distribution, larger mutation rates and smaller selection coefficients needed to be investigated to find the best-fitting pair: we scanned mutation rates from $\log_{10}(\mu) = -8$ to $\log_{10}(\mu) = -3$, and selection coefficients from 0.002 to 0.1, with increments of 0.002. Additionally, in the exponential distribution, the majority of mutations had a fitness value that was too small to contribute to the competition between the two cell populations (YFP and CFP). However, increasing the mutation rate would increase the probability that mutations with larger fitness value were obtained. We found that for mutation rates higher than 10^{-4} per genome per cell division, the computational time constraints were substantial; thus we excluded all mutations that had an s value smaller than 10% of the average fitness effect s . Support for this choice was provided by the finding that for large mutation rates (that is, $N\mu > 1$), the mutations chosen from an exponential distribution that eventually reach fixation within the population had a larger fitness effect than the distribution average (ref. 20 and Fig. 3). The above criteria for large mutation rates were within our experimental regime ($N\mu = 10^6 \times 10^{-4} > 1$). Furthermore, in Extended Data Fig. 3e we present simulation results of the average deviation from equal percentages of CFP and YFP for a given mutation rate and fitness effect drawn from an exponential distribution. We found that the results were robust to including or excluding those mutations with fitness effects smaller than the distribution average. Thus, for high mutation rates ($N\mu > 1$), we could exclude weak mutations, as was shown in ref. 20.

The fitting of the empirical experiments and the simulations was done as described above. The results of the best-fit values are presented in Extended Data Fig. 2a–d.

The equivalence principle model. As a complementary approach we evaluated the equivalence principle model developed in ref. 17 to analyse a similar competition experiment in *Escherichia coli*. This study concluded that the mutations that eventually reach fixation in large microbial populations have a very narrow range of fitness values. Very weak mutations are unlikely to lead to a takeover of one population (CFP or YFP) by the other, whereas very strong mutations are very infrequent. On the basis of this prediction, the competition experiments were described by assuming that all beneficial mutations have exactly the same fitness value; that is, assuming that the distribution of fitness effects is a delta function. On the basis of this assumption, the authors developed a method for inferring the two key parameters of the dynamics of the adaptation process: the mutation rate of a beneficial mutation and the single fitness effect value. Note that the assumption of a single fitness value for beneficial mutations was shown to be true only if the mutation rate is very high²⁰, which cannot be known *a priori*. This is why the branching evolution model was implemented first and then compared with the equivalence principle model. Note also that the equivalence principle model analyses only the initial CFP:YFP divergence phase, rather than using the entire data set as in the branching evolution model. Although this excludes some data, it has the advantage of not necessitating the assumption that the distribution of fitness effects is constant⁴⁰.

We followed the procedure outlined in ref. 17, including the following assumptions and modelling steps (see the Supplementary Information in ref. 17 for a more complete description).

1. Rather than the computationally intensive simulation of growth and dilution, we used the effective population size⁴¹. The effective population size is $N_e = N_0 \log(r)T$, where N_0 is the initial population size, r is the growth rate of the population, and T is the number of generations between dilutions. In addition to reducing computation time, the use of the effective population size allows for an analytical approximation of certain quantities, such as the probability of escaping drift.

2. Even large mutations may be eliminated by drift, and only those that escape drift can contribute to the competition between the two labelled cell populations. The probability of emergence of a new mutation that will escape drift can be calculated analytically for a fixed population, making it possible to generate in simulations only those mutations that escaped drift, and contribute to the competitiveness of the population. Under the above simplifying approximations of (1) a fixed population size and (2) a delta distribution of fitness effects, the probability that a mutation will emerge and escape drift (P_{escape}) was calculated analytically, as a function of the mutation rate, μ (the rate for a new mutation to appear in a single cell division), the selection coefficient, s , and the effective population size, N_e .

3. Given that P_{escape} can be calculated, simulations can be performed rapidly by generating the time in the experiment when such mutations occur, rather than randomly assigning mutations at each cell division.

We applied the equivalence principle model as follows. To compare the experiments with the simulations, as in ref. 17, we used time and initial slope of CFP:YFP ratio divergence as summary statistics for the divergence curves. Each CFP:YFP divergence experiment (one single well from a 96-well plate) was fit to an exponential growth model, with slope α that starts at a given time τ , using the expression $-\log_{10}(1 + 0.5\exp(\alpha(t - \tau)))$. An exponential model is a good description of the initial divergence phase. The end of the exponential growth phase was defined by the time that had the maximum likelihood to be described by exponential growth (see the Supplementary Information in ref. 17 for more details). The initial CFP:YFP bias that we detected in some 2N and 4N populations was included in cases where it occurred.

Each divergence experiment is thus described by its values τ and α , whether it is a real experiment or a simulated one. Because a single divergence experiment is subject to significant stochasticity, we compared the empirical experiments and the simulations by combining all the wells of each ploidy type. This was done by collecting all of the α values of a given experiment or a given set of simulations (that is, those simulations that we generated by the same pair of μ and s values) into a distribution of α values. The experimental distribution was compared with the distribution of the simulations by a Kolmogorov–Smirnov test between the two. The same was done for the τ values. For each pair of s and μ values, we calculated the sum of the Kolmogorov–Smirnov test of its α distribution and the Kolmogorov–Smirnov test of its τ distribution. The pair of values (μ, s) that had the smallest sum was declared the best fit (Extended Data Fig. 2a). This was done independently for each ploidy type (Extended Data Fig. 2e).

Modelling results. The fitting of the empirical experiments and the simulations was done as described above. The results of the best-fit values from the branching evolution model and the equivalence principle model are presented in Extended Data Fig. 2a. To generate the error range for each distribution and ploidy type, we used the parametric bootstrapping method⁴². For each estimated set of values, we generated 1,000 simulated data sets and compared those with the empirical data sets (data set sizes were 172, 264, and 265 independent deviation experiments for the haploids, diploids, and tetraploids, respectively). We then inferred the best-fit values of those data sets. The 95% confidence intervals of μ and s from those 1,000 data sets were defined as the error ranges (Extended Data Fig. 2a–e).

We infer from these results that the tetraploids have a higher mutation rate and that these mutations have, on average, a stronger fitness effect compared with haploids or diploids. The trend for tetraploids to have higher μ and s values occurs independently of the different assumed distributions of fitness effects. However, the different distributions lead to significantly different absolute values. This is expected from the characteristic shape of these distributions. To illustrate why this result is expected, we show a schematic diagram of the three distributions of fitness effects (Extended Data Fig. 2f).

The mutations that mainly govern the adaptation process are hypothesized to come from a relatively narrow range of fitness effect values that are sufficiently strong, but not extremely rare (Extended Data Fig. 2f, double arrow region from the idealized narrow Gaussian curve). The three assumed distributions that we used in our modelling approximate the true distribution but with the following differences. The delta distribution is located in the centre of the double arrow region. Thus, by definition, every mutation from the delta distribution will have a fitness effect that is strong enough to have a significant probability of promoting adaptation. If every mutation has the possibility of contributing, then a lower mutation rate suffices for adaptation. By contrast, the exponential distribution is dominated by small-effect or near-neutral mutations, with a relatively low fraction of mutations near the value for the delta distribution. Thus, the assumption of an exponential

distribution is accompanied by a requirement for a compensatory higher mutation rate to achieve numbers of equivalently beneficial mutations. By the same line of reasoning, the uniform distribution necessitates a rate of beneficial mutations that is intermediate between the exponential and delta distributions. Indeed, our results match all of these expectations (Extended Data Fig. 2a–d). A similar difference in the estimation of μ was also observed in another recent study that modelled this value based on these three assumptions about the distribution of beneficial mutations⁴³.

In terms of the fitness effect values of the different distributions, the uniform distribution is governed by its upper limit, which, by definition, is the strongest mutation allowed. Therefore, it is expected that the uniform distribution mean, which is half of the upper limit, will be half of the fitness effect value of the delta function. Again, this is what we observed (Extended Data Fig. 2a–d). The exponential distribution mean, is much smaller than the strongest mutation that can be generated by the exponential distribution, and the best-fit parameters are expected to be much smaller than the other two distributions, which is also observed (Extended Data Fig. 2a–d). Furthermore, because the exponential distribution has no upper bound on the fitness effects, a larger mutation rate can lead to the emergence of mutations with much larger fitness effects. In this way, under the exponential distribution assumption, the mutation rate affects the range of possible fitness effects, which results in a larger error range for the exponential distribution compared with the other distributions (Extended Data Fig. 2a, brackets).

In summary, despite expected differences in absolute values, we reach the same overall conclusions with either the branching evolution model with varied assumptions about the distribution of beneficial mutations or the equivalence principle model¹⁷.

Computer code availability. The computer code is available as a Supplementary Software file. The code was compiled by g++ (version 4.2.1) and was tested on Unix (CentOS5 operating system) and Mac (OS X version 10.9.2) machines.

Plasmid construction. All plasmids used in this study are listed in Extended Data Table 1. To construct plasmids for the inducible expression of either *CFP* or *YFP*, the galactose-inducible *GAL1* promoter was subcloned into the YFP plasmid PB1500 and the CFP plasmid PB2452. These plasmids were derived from the green fluorescent protein (GFP) tagging plasmid generated in ref. 44. Both plasmids contained the *ADH* gene terminator (*tADH*) after the *YFP* or *CFP* gene and the sequence of the *SpHIS5* gene of *Schizosaccharomyces pombe* as a selectable marker. Plasmids PB1500 and PB2452 were digested with BamHI and PacI to introduce the pGAL promoter, 461 base pairs (bp) upstream of the start codon of *GAL1* (ref. 45), which was amplified using the primers pGAL1 BamHI 5' (5'-ACGGATCCCCGGGTGAA GTACGGATTAGAAGCCGCCGAG-3') and pGAL1 PacI 3' (5'-CGTTAATTA ATATAGTTTTTCTCTTGACGTAAAG-3'). Site-directed mutagenesis (Quick Change Mutagenesis Kit, Stratagene) was used to introduce an ATG translation start codon to the *YFP* and *CFP* genes (using the GAPATGpFA6 primer, 5'-CAA TCAATCAATCAATCATCACATAAATTAATTAATGAGTAAAGGAGAAG AACTTTTCACTGGAGTTGTC-3'). The resulting plasmids PB2694 and PB2697 contained the cassette pGAL1-*CFP*-*tADH*-*SpHIS5* and pGAL1-*YFP*-*tADH*-*SpHIS5*, respectively.

PB2314 was used to delete the *MAT* locus as previously described¹². PB1308 was used to perform a *URA3* to *TRP1* marker swap, as previously described⁴⁶. PB1640 (hphMX4, ref. 47) was used for PCR-mediated deletion of *STE4*. PB2647 (*STE4*-*LEU2*) was used to restore mating competency and was constructed by amplifying the *STE4* gene with primers *STE4* P BamHI 5' (5'-CCGGATTCTTGTAGCCCTG TTAGTTTACC-3') and *STE4* T BamHI 3' (5'-CCGATTCCAATACATAAG GACGAGCCAGTG-3'), and cloning it into pRS315. PB2649 (*STE4* *URA3* *CEN* *MAT α*) was also used to restore mating competency, and was constructed by sub cloning the *STE4* fragment from PB2647 (digested with SmaI and NotI) into PB2577 (*MAT α* *URA3* *CEN*, digested with SmaI and NotI).

Yeast strain construction. All *S. cerevisiae* strains used in this study were isogenic to PY3295 (BY4741, S288c genetic background *MATa his3A leu2A met15A ura3A*) and are listed in Extended Data Table 1. The strategy used to generate the isogenic ploidy series is illustrated in Extended Data Fig. 1 and genotypes of key intermediates are indicated. The CFP and YFP ancestors were derived from the haploid strain PY5997 (*matA::pSTE5-ura3::TRP1*, *ste4A::Hygro^R*, *trp1::Nat^R*, strain construction details available upon request). Isogenic strains with either the *CFP* or *YFP* cassettes at the *TRP1* locus (chromosome IV) were generated (PY5998 and PY5999) as follows: the pGAL1-*CFP*-*tADH*-*SpHIS5* or pGAL1-*YFP*-*tADH*-*SpHIS5* cassette was PCR amplified from plasmid pB2694 (*CFP*) or pB2697 (*YFP*), respectively, with primers delTRPGFP5' (5'-TATTGACAGCTGAGTATACGTGATTAAG CACACAAAGGCAGCTTGAGTGCAGGTCGACGATCCCGGG-3') and delTRPGFP3' (5'-GAACGTGCACTGAGTAGTATGTTGCAGTCTTTTGGA ATACGAGTCAATTCGAGCTCGTTAAAC-3') and transformed into PY5997 at the *TRP1* locus. The haploid ancestor strains expressing *CFP* (PY5998) or *YFP* (PY5999) were confirmed by PCR and fluorescence microscopy. The haploid ancestors

were modified to become mating competent by transformation with plasmids PB2647 (*LEU2*-*STE4*) and PB2649 (*URA3*-*STE4*-*MAT α*). Diploid zygotes were selected on -Ura -Leu plates, and then colony purified on YPD plates to allow plasmid loss. Diploid chromosome content was confirmed by flow cytometry and aCGH, and strains PY6006, PY6008, PY6014, and PY6022 were selected. The diploid ancestors were made mating competent by transformation with plasmids PB2647 and PB2649. Tetraploid zygotes were pulled onto YPD plates using a micromanipulator, and after 2 days growth at 30 °C the ploidy of each zygote was determined by flow cytometry and aCGH.

The *SNF3*-*G439E* mutation was constructed in the haploid YFP strain background (PY5999) using the pCORE counter-selectable reporter system⁴⁸, a gift from M. Resnick. Primers *SNF3*_pCORE_KAN (5'-TGTTGGGGGTGTTATCATGA CTATAGCCCACTTTATGTGGCCATTGTTGGGAGTCGTTTCGACAC TGG-3') and *SNF3*_pCORE_URA (5'-TATAAATGCTATCATAAAGTTTTCG GCGCTACAGTCTTTAAGGAACACTCCTTACCATTAAAGTTGATC-3') were designed to integrate the *CORE* sequence at the *SNF3* locus; PCR amplification and transformation procedures were followed as detailed previously⁴⁹. Sanger sequencing was used to identify clones with the desired mutation (chromosome IV: 112,896 G>A). Diploid *SNF3*-*G439E* mutants (heterozygous *SNF3*-*G439E*/*SNF3* and homozygous *SNF3*-*G439E*/*SNF3*-*G439E* clones) were constructed by mating after introduction of plasmids to confer mating competence (PB2649 or PB2647), as described above for the construction of the CFP- and YFP-marked strains. An analogous strategy was used to generate tetraploid *SNF3*-*G439E* strains (heterozygous *SNF3*-*G439E*/*SNF3*/*SNF3*/*SNF3*).

The chromosome XIII aneuploid strain series was constructed in the S288c background from the diploid strain PY7295 (RL4737) and the diploid PY7296 (RL4888), which is trisomic for chromosome XIII⁵⁰. PY7296 was isolated from a triploid meiosis and a minimal number of cell divisions⁵⁰. We confirmed the chromosome XIII trisomy by aCGH (Extended Data Fig. 9). We generated tetraploid clones by mating PY7295 to PY7296, with changes in mating-type accomplished as described previously¹². Tetraploid clones were isolated on selective media and analysed by flow cytometry and aCGH (representative clones Extended Data Fig. 9). Additional details for all yeast strain constructions are available upon request.

Relative fitness assays. Competitive fitness assays were performed using single colony isolates from the evolved populations and a common ancestor. One single colony was isolated from frozen stocks of each well of the evolution experiments (1N(A), 1N(B), 2N(A), 2N(B), 4N(A), 4N(B), 4N(C)) at generation 250. The evolved clones were cultured for 24 h in 500 μ l of Synthetic Complete + 2% raffinose, diluted into fresh medium, and competed with the ancestor expressing the complementary fluorescent protein. Competitions were initially performed using approximately the same number of cells from the ancestor and the evolved clone, but because the evolved clones grew significantly faster than the ancestor strains, the competitions were repeated using approximately five times more ancestor cells than evolved clone cells, with an initial population size of 1×10^5 . Serial dilutions were performed each day and the YFP:CFP ratio was determined by flow cytometry, yielding an estimate of the number of evolved (N_t) cells relative to the ancestors (N_0) as a function of time. The data were analysed in Matlab using a custom script that performed a linear least-squares fit of $\log(N_t/N_0)$ over multiple dilution cycles. The fitness relative to the ancestor is defined as $s = d/dt [\log_2(N_t/N_0)]$, where t is measured in days¹⁷.

Flow cytometry analysis of DNA content. Cells were prepared for propidium iodide staining as described but with modifications to optimize preparation of samples in 96-well plates. Thirty thousand cells were analysed using a BD LSRII High Throughput Sampler. FlowJo cell cycle analysis used the Dean-Jett-Fox model to estimate the mean G1 and G2 fluorescence peaks of each strain. Control parental 1N, 2N, 3N, and 4N strains were analysed in triplicate with the evolved strains.

Microarray aCGH. Fluorescently labelled DNA was prepared for CGH as described previously⁵¹. Genomic DNA from all experimental strains was compared with the same pool of genomic DNA from the ancestral strain background PY3295 (BY4741, Research Genetics). Agilent yeast DNA 4 \times 44K microarrays (ChIP-on-chip Kit) were used for the hybridization according to the manufacturer's instructions (Agilent Technologies) with several modifications (M. Dunham online protocols, <http://dunham.gs.washington.edu/protocols.shtml>). Briefly, 2.0 μ g of HaeIII-digested (New England Biolabs) genomic DNA was labelled with 2.1 μ l of Cy3 or Cy5 (CyDye-Cy3-dUTP or CyDye-Cy5-dUTP, Amersham GE Healthcare). Three hundred nanograms of Cy3-labelled DNA (experimental strains) was mixed with 300 ng of Cy5-labelled DNA (control DNA) and the volume was brought to 44 μ l with nuclease-free water. Blocking buffer and hybridization buffer 2 \times HiRPM (Agilent Technologies) were added, and 100 μ l was applied to each sub-array; the microarray was hybridized at 65 °C for 17 h and then washed, scanned, and analysed according to the manufacturer's instructions. Agilent Feature Extraction data were converted from \log_{10} ratios to \log_2 ratios and plotted using Treeview⁵² and a custom Matlab script.

A log₂ ratio of zero (baseline) indicates no difference in DNA copy number between reference and experimental samples^{14,51}.

Quantitative PCR. All qPCRs were performed on an Applied Biosystems ViiA-7 real-time PCR machine in 96-well format with Power SYBR Green PCR Master Mix (Applied Biosystems) and three technical replicates. *HXT6/7* gene copy number was determined relative to the ancestor as previously described⁵³. Genomic DNA was isolated and RNase treated from 30 clones with the highest fitness from each haploid, diploid, and tetraploid experiment. *HXT6/7* on chromosome IVR was amplified using forward primer 5'-GATTATTGCTGGTCCGATCC-3' and reverse primer 5'-GAGTAATCGCCAATGGGTCT-3'; the control loci *UBP1* on chromosome IVL was amplified using forward primer 5'-GCGCTCTGTCATTGTTTAC T-3' and reverse primer 5'-GACTTTTCAGCTTCGTCCACAA-3'. Raw *HXT6/7* values were normalized to *UBP1* for each clone and then normalized to the ancestor ($\Delta\Delta Ct$). We found the amplification in 3% of the 1N-, 30% of the 2N-, and 43% of the 4N-evolved clones ($n = 30$). This significant bias for *HXT6/7* amplification in the 2N and 4N populations (t -test, $P = 0.005$ and $P = 1 \times 10^{-4}$, respectively) might have been due to mutations in 1N cells that prevented the acquisition of the *HXT6/7* amplification because of negative epistasis²⁴.

SNF3 gain-of-function mutations were previously shown to increase expression of *HXT4* (ref. 25). Therefore we analysed *HXT4* gene expression levels in diploid-evolved clone 2N_233 (carrying the *SNF3-G439E* mutation), relative to the diploid ancestor (PY6006). Strains were grown up from -80°C stocks overnight in 5 ml Synthetic Complete + 2% raffinose and then diluted into fresh 25 ml Synthetic Complete + 2% raffinose. Cells were cultured at early log phase and RNA was extracted using an RNeasy Mini Kit (Qiagen). Complementary DNA was prepared using a SuperScript III First-Strand synthesis system (Life Technologies). *HXT4* was amplified using primers 5'-TAAGTCAGCGCAGACGATCCA-3' and 5'-TTCACC CCAGGAGGCATTACCA-3'; *ACT1* was amplified using primers 5'-ACGTCGC CTTGGACTTCGAACA-3' and 5'-TGGAACAAAGCTTCTGGGGCTC-3'. Raw qPCR values were normalized to *ACT1* levels and then normalized to the 2N ancestor (PY6006). Relative to the ancestor, the clone bearing *SNF3-G439E* had eightfold higher *HXT4* expression in raffinose medium.

Overview of WGS. We performed WGS and identified *de novo* variants for 74 evolved clones and two ancestors. Initially six evolved clones and one tetraploid ancestor were sequenced on ABI's SOLiD 4 platform. Subsequent sequencing of 68 evolved clones and the haploid ancestor was performed on an Illumina HiSeq 2500. The specifics of the analysis pipeline for each sequencing platform are provided below. Regardless of the underlying platform, the overall analysis strategy was as follows. Briefly, the raw reads underwent quality analysis and barcode/adaptor removal. High-quality reads were mapped to the *S. cerevisiae* reference genome (downloaded June 2010). Reads containing PCR-based artefacts were removed and alignments underwent local realignment around insertions and deletions (indels), resulting in the highest quality alignment. SNPs and indels were called and combined across the evolved strains and within the parental strains to identify a set of variants in the strain background relative to the reference. Each evolved strain (all 74) was individually compared with the parental set to identify the set of potential *de novo* variants. These evolved strain calls were filtered by quality metrics and manually inspected. All variants of moderate or poor quality as well as a few good quality variants were analysed by Sanger sequencing. Chromosomal aneuploidy was inferred from changes in read depth using a windowing approach.

Per cell, the evolved tetraploids had more mutations than haploids or diploids. However, per haploid genome, the evolved tetraploids on average accumulated a similar number of mutations (1.50 average SNPs per haploid genome, on the basis of final evolved ploidy) as the evolved diploids (1.44 average SNPs per haploid genome) and fewer mutations than the evolved haploids (2.05 average SNPs per haploid genome); however, neither comparison was significant. It is likely that the number of mutations in the evolved tetraploids was underestimated because of the high rate of chromosome loss in these strains. Interestingly, there was a higher average for the tetraploid-evolved clones that became $\sim 2\text{N}$ in ploidy (2.00 average SNPs per haploid genome) than the diploid-evolved clones (1.44 average SNPs per haploid genome), suggesting that near-diploid cells that underwent a tetraploid intermediate might have acquired more mutations than cells that remained diploid throughout the experiment. However, despite the trend, this effect did not reach statistical significance ($P = 0.22$).

Illumina sequencing (2 × 100). Library preparation. Clones selected for WGS were cultured overnight from -80°C stocks in 2 ml Synthetic Complete + 2% raffinose medium. Genomic DNA was isolated using phenol–chloroform–isoamylalcohol (24:25:1) and bead beating. Libraries were prepared as described⁵⁴. Briefly, DNA was sheared with Diagenode Bioruptor (UCD-200) to a median size of 300–500 bp, end-repair was performed with an NEB Next End repair kit (NEB E6050L), and fragments were A-tailed with Klenow fragment (M0212L). Custom adaptors with in-line barcodes were ligated overnight. Adaptor ligated fragments were size selected on 1% TBE agarose gel stained with Sybr Gold (Invitrogen S-11494) for fragments between

400 and 600 bp and isolated using Qiagen Gel Extraction Kit (28706). Libraries were amplified for 12 cycles with Illumina PE PCR primers 1.0 and 2.0. Libraries were pooled and underwent additional size selection for fragments of 400–600 bp.

Raw data. The genomes were sequenced on an Illumina HiSeq 2500 at the University of Colorado at Denver Next Generation Sequencing Facility. The data, which had an inline barcode, were de-multiplexed by the sequencing facility into individual sample R1/R2 files—one file for each read in the pair. The barcodes were removed before mapping using Fastx_trimmer (version 0.0.13.2, http://hannonlab.cshl.edu/fastx_toolkit/). Read trimming from the 5' end of the R2 reads was performed on a sample-specific manner trimming anywhere from 0 to 28 base pairs using in-house script and Fastx_trimmer.

Mapping. Reads were mapped to the *S. cerevisiae* reference sequence for the laboratory yeast strain S288c reference genome (*S. cerevisiae* genome obtained 28 July 2010 from the Saccharomyces Genome Database FTP site: http://downloads.yeastgenome.org/sequence/S288C_reference/genome_releases/; actual genome: http://downloads.yeastgenome.org/sequence/S288C_reference/genome_releases/S288C_reference_genome_R63-1-1_20100105.tgz). The reads were mapped using the Bowtie2 (version 2.0.2)⁵⁵ local alignment strategy, allowing for multiple mapping, and setting the following options: `-very-sensitive-local -I 180 -X 1,000 -score-min G,70,8`. The mapped reads then underwent file format conversion into the binary format for downstream analysis using Samtools view, sort, and index (version 0.1.18)⁵⁶.

Alignment tailoring. Post-alignment to the genome, duplicate pairs resulting from PCR overamplification were removed using Samtools rmdup, eliminating 1–5% of the paired reads. The reads were re-aligned over potential indel sites using the Genome Analysis Toolkit RealignerTargetCreator and IndelRealigner (version 2.4-9)^{57,58}.

Variant calling and refinement. Variant calling was performed on the tailored read mappings using Genome Analysis Toolkit UnifiedGenotyper (version 2.4-9)^{57,58}. For the haploids and diploids, SNPs were called using default parameters; for the higher ploidy strains, the ploidy option was increased to $5n$, which allowed identification of mutations at allelic frequencies down to 5% alternative allele representation. Variant lists were combined on the basis of ploidy-type using Genome Analysis Toolkit CombineVariants. SNPs and short indels were compared with the parental set of mutations using in-house scripts to generate a set of non-Parental mutations. These mutations were filtered for alternative allele support and allelic frequency (more than two reads supporting an alternative allele for coverage $10\times$ to $20\times$, and more than four reads supporting an alternative allele for coverage more than $20\times$). The filtered mutations were manually inspected using the Integrative Genome Viewer (version 2.1.19)⁵⁹ to refine the set and further remove mapping artefacts such as strand representation bias, regional mapping quality issues from non-unique mapping, and artefacts of homopolymer and simple repeat alignments. We Sanger sequenced variants with low read support (fewer than five reads supporting an alternative allele), as well as a subset of the other medium- and high-confidence variants. The final set of evolved variants discovered was annotated versus the gene file specific to the genome using an in-house script.

Chromosomal CNV identification. Identification of chromosomal CNVs used HTSeq (version 0.6.1)⁶⁰ in conjunction with custom scripts. HTSeq performs coverage estimations on a per-gene basis, and the custom scripts provided normalized log₂ fold change between each sample and the parental haploid strain. Estimates on chromosomal copy number were inferred using the median value for the log₂ fold change on a chromosome-by-chromosome basis. We implemented the Cochran–Armitage test to determine whether chromosome XIII had a trend for higher copy number, relative to the copy number observed for all chromosomes in the tetraploid-evolved clones (Fig. 2d and Supplementary Table 2). This trend analysis is similar to a χ^2 test, but tests whether there is a significant trend or direction to the observed data set (chromosome XIII copy number).

Sequencing quality assessment. Because our sequencing was highly multiplexed, quality assessment on the sequencing data was necessary to eliminate strains without adequate genome coverage. For the haploids and diploids, we determined the adequate depth of coverage to recover mutations in two ways. First, we took the set of 'strain-background' mutations, which were identified by filtering the parental variant calls for a conservative, high-quality (quality > 100), homozygous set of locations. Each strain was then queried for the ability to recapitulate these variants, reporting a percentage overlap between each strain's variant calls and the set of background variants. Any strain with less than 97% of the background mutations was dropped from consideration. Additionally, we examined the impact of subsampling down to various depths to investigate the impact of lower coverage on recovering variants. This was done using Picard's Downsampler.jar (version 1.72, <http://broadinstitute.github.io/picard/>) on two higher-coverage diploid strains to randomly down-sample the coverage to $100\times$, $50\times$, $25\times$, and $10\times$ coverage. We examined the SNP call overlap and found that for the strain-unique SNPs they could be captured even at a level of $10\times$ coverage. Using this information, we set minimum coverage requirements for each strain on a genome-wide scale to eliminate strains

without adequate genomic representation. We analysed depth of coverage on all of the mapped data with BEDTools genomeCoverageBed (version 2.16.2)⁶¹. The per-base coverage was then analysed using an in-house script to produce statistics on minimum coverage per allele, average coverage, etc.

SOLiD sequencing (2×50). Library preparation. We performed a pilot experiment on seven strains using SOLiD paired-end sequencing (Supplementary Table 1). Clones selected for SOLiD sequencing were cultured overnight from −80 °C stocks in 4 ml Synthetic Complete + 2% raffinose medium. Genomic DNA was isolated using QIAGEN Genomic-Tip 100 according to the manufacturer's instructions. SOLiD library preparation and sequencing was performed by the Molecular Biology Core Facility at the Dana-Farber Cancer Institute according to the manufacturer's instructions (Applied Biosystems, Life Technologies).

Mapping. The sequencing reads were mapped to the *S. cerevisiae* reference genome (see 'Mapping' in the Illumina section above) using multiple different mapping software including BWA (version 0.5.9)⁶², NovoAlignCS (version 1.01.05)⁶³, Bfast (version 0.6.5a)⁶⁴, and BowtieCS (version 0.12.7)⁶⁵. BowtieCS and BWA were used in the downstream variant calling and copy number changes, while NovoAlignCS and Bfast served as added support in manual inspection of variants.

Alignment tailoring. After mapping, the reads were post-processed for local realignment using SRMA (version 0.1.15)⁶⁶ and Samtools BAQ (version 0.1.18)⁵⁶.

Variant calling and refinement. SNPs, small insertions and deletions (indels) were called from the post-processed reads using Samtools Mpileup (version 0.1.18)⁵⁶, VARIid (version 1.0.7f)⁶⁷, and Freebayes (version 0.8.9, <http://bioinformatics.bc.edu/marthlab/FreeBayes>). Samtools and VARIid variant calls were used to identify the strain background (parental variants relative to the reference). These variants were filtered on the basis of reads supporting the allele in both directions, quality score of the call, and adequate read coverage over the call. Once filtered, all of the variant calls for the evolved strains were merged and compared with the parent. Variations were verified by manual inspection followed by Sanger validation for both a set of randomly sampled loci and regions of disagreement between different combinations of the mapping software and the variant callers (that is, dinucleotide SNPs and multiple indels within a single read). The resulting set was later used for identification of strain-unique variants in the evolved strains.

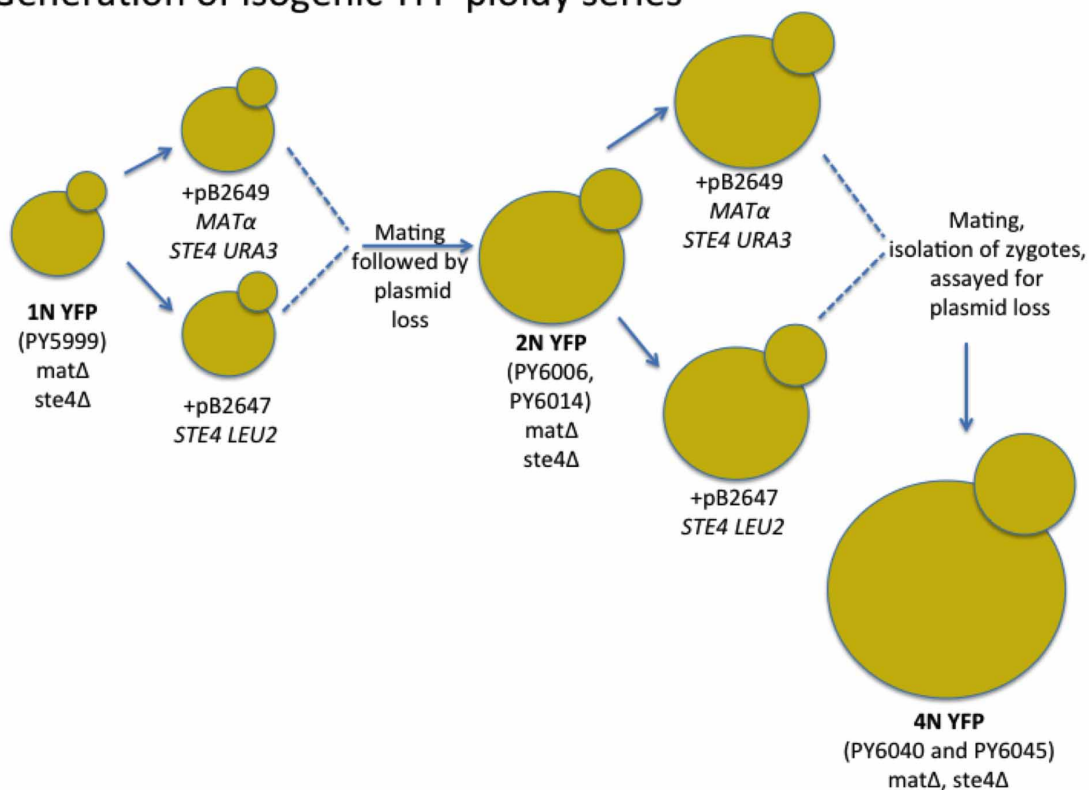
To identify strain-unique variants, Freebayes, a variant caller capable of higher-ploidy (ploidy > 2N), was used. Freebayes has the ability to set the assumed ploidy over a genomic region to adjust the expected distribution for allelic frequency. The assumed ploidy was determined using aCGH as well as the copy number changes. The Freebayes called variants on each evolved progeny were then cross-referenced with the parental variants to produce strain-unique variants (Supplementary Table 1). These variants were then manually examined in Integrative Genomics Viewer⁵⁹ and validated by Sanger sequencing. PCR amplification and Sanger sequencing of approximately 200 bp on either side of the sequence variants used DNA from the evolved clone and the ancestor.

Chromosomal CNVs. Copy number changes, first identified in the aCGH data, were confirmed in the genomic DNA sequencing using BedTools genomeCoverageBed⁶¹ in combination with custom in-house scripts and DESeq (version 1.10.1)⁶⁸. Briefly, the normalized genomic copy number of all annotated genes in each strain was compared back with the parent. These comparisons were then plotted using an in-house script (see Chromosomal CNV identification above).

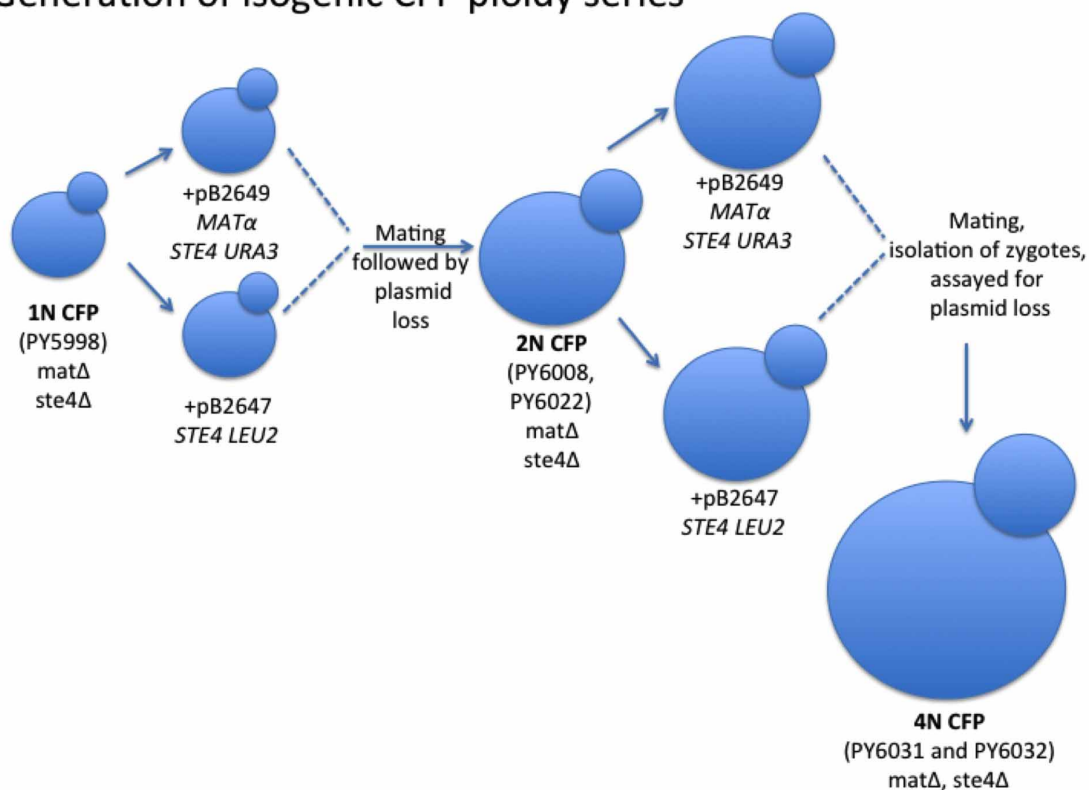
31. Durrett, R., Foo, J., Leder, K., Mayberry, J. & Michor, F. Intratumor heterogeneity in evolutionary models of tumor progression. *Genetics* **188**, 461–477 (2011).
32. Zhu, Y. O., Siegal, M. L., Hall, D. W. & Petrov, D. A. Precise estimates of mutation rate and spectrum in yeast. *Proc. Natl Acad. Sci. USA* **111**, E2310–E2318 (2014).
33. Barbera, M. A. & Petes, T. D. Selection and analysis of spontaneous reciprocal mitotic cross-overs in *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **103**, 12819–12824 (2006).
34. Rouzine, I. M., Wakeley, J. & Coffin, J. M. The solitary wave of asexual evolution. *Proc. Natl Acad. Sci. USA* **100**, 587–592 (2003).
35. Desai, M. M., Fisher, D. S. & Murray, A. W. The speed of evolution and maintenance of variation in asexual populations. *Curr. Biol.* **17**, 385–394 (2007).
36. Fogle, C. A., Nagle, J. L. & Desai, M. M. Clonal interference, multiple mutations and adaptation in large asexual populations. *Genetics* **180**, 2163–2173 (2008).

37. Vetterling, W. T. *Numerical Recipes Example Book* 2nd edn, Ch. 7 (Cambridge Univ. Press, 1992).
38. Wakeley, J. *Coalescent Theory: An Introduction* (Roberts, 2009).
39. Moura de Sousa, J. A., Campos, P. R. & Gordo, I. An ABC method for estimating the rate and distribution of effects of beneficial mutations. *Genome Biol. Evol.* **5**, 794–806 (2013).
40. Goyal, S. et al. Dynamic mutation-selection balance as an evolutionary attractor. *Genetics* **191**, 1309–1319 (2012).
41. Ewens, W. J. *Mathematical Population Genetics* 2nd edn (Springer, 2004).
42. Efron, B. & Tibshirani, R. *An Introduction to the Bootstrap* (Chapman & Hall, 1993).
43. Frenkel, E. M., Good, B. H. & Desai, M. M. The fates of mutant lineages and the distribution of fitness effects of beneficial mutations in laboratory budding yeast populations. *Genetics* **196**, 1217–1226 (2014).
44. Longtine, M. S. et al. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**, 953–961 (1998).
45. Mumberg, D., Muller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119–122 (1995).
46. Cross, F. R. 'Marker swap' plasmids: convenient tools for budding yeast molecular genetics. *Yeast* **13**, 647–653 (1997).
47. Goldstein, A. L. & McCusker, J. H. Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* **15**, 1541–1553 (1999).
48. Storici, F., Lewis, L. K. & Resnick, M. A. In vivo site-directed mutagenesis using oligonucleotides. *Nature Biotechnol.* **19**, 773–776 (2001).
49. Storici, F. & Resnick, M. A. The *delitto perfetto* approach to in vivo site-directed mutagenesis and chromosome rearrangements with synthetic oligonucleotides in yeast. *Methods Enzymol.* **409**, 329–345 (2006).
50. Pavelka, N. et al. Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* **468**, 321–325 (2010).
51. Selmecki, A., Bergmann, S. & Berman, J. Comparative genome hybridization reveals widespread aneuploidy in *Candida albicans* laboratory strains. *Mol. Microbiol.* **55**, 1553–1565 (2005).
52. Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics* **20**, 3246–3248 (2004).
53. Wenger, J. W. et al. Hunger artists: yeast adapted to carbon limitation show trade-offs under carbon sufficiency. *PLoS Genet.* **7**, e1002202 (2011).
54. Hittinger, C. T. et al. Remarkably ancient balanced polymorphisms in a multi-locus gene network. *Nature* **464**, 54–58 (2010).
55. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).
56. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
57. McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
58. DePristo, M. A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genet.* **43**, 491–498 (2011).
59. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
60. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
61. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
62. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
63. Novocraft. Novocraft short read alignment package. <http://www.novocraft.com> (2009).
64. Homer, N., Merriman, B. & Nelson, S. F. BFAST: an alignment tool for large scale genome resequencing. *PLoS ONE* **4**, e7767 (2009).
65. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
66. Homer, N. & Nelson, S. F. Improved variant discovery through local re-alignment of short-read next-generation sequencing data using SRMA. *Genome Biol.* **11**, R99 (2010).
67. Dalca, A. V., Rumble, S. M., Levy, S. & Brudno, M. VARIid: a variation detection framework for color-space and letter-space platforms. *Bioinformatics* **26**, i343–i349 (2010).
68. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

Generation of isogenic YFP ploidy series



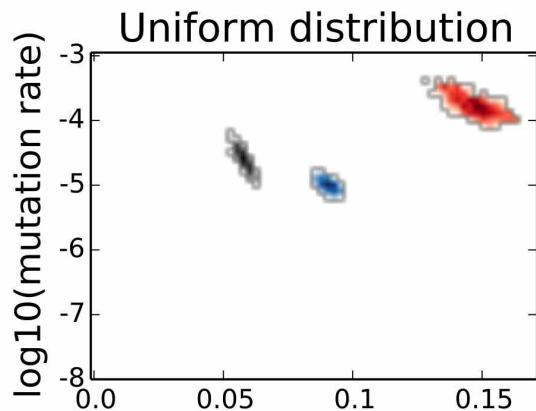
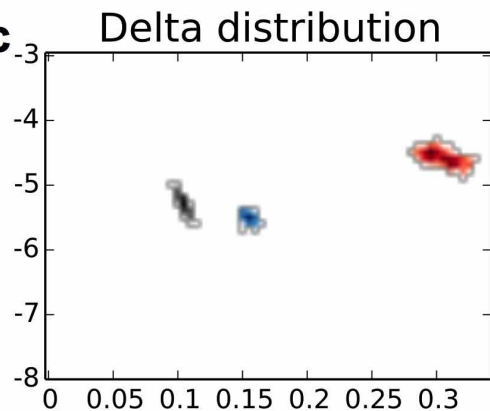
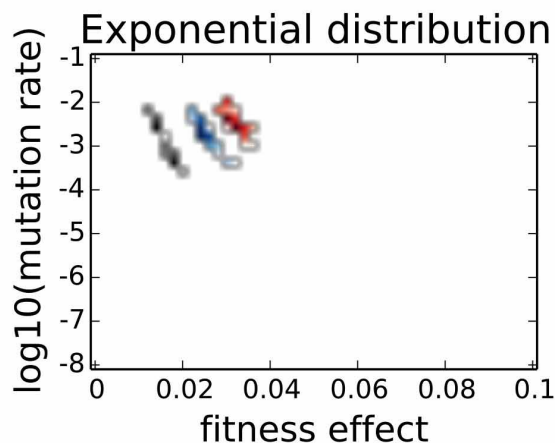
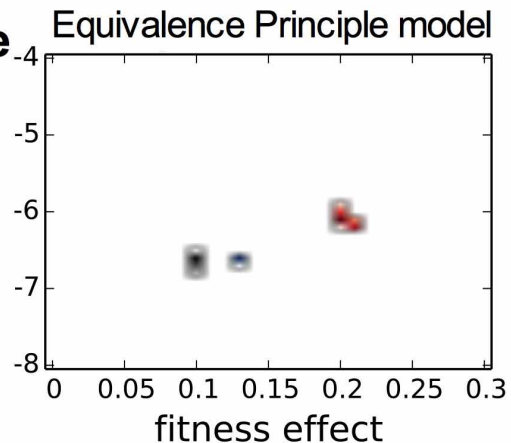
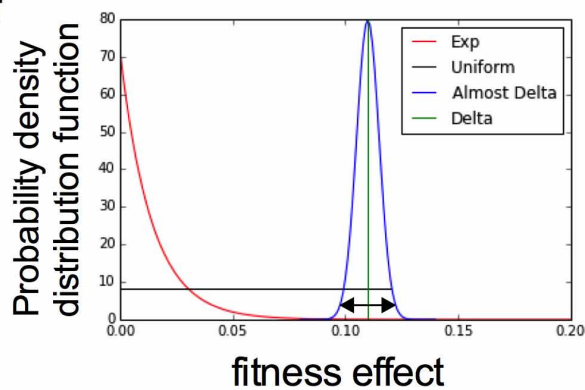
Generation of isogenic CFP ploidy series



Extended Data Figure 1 | Schematic representation of the construction of isogenic haploid, diploid, and tetraploid strains used in this study. Relevant strain numbers are indicated for the CFP- and YFP-containing ancestors.

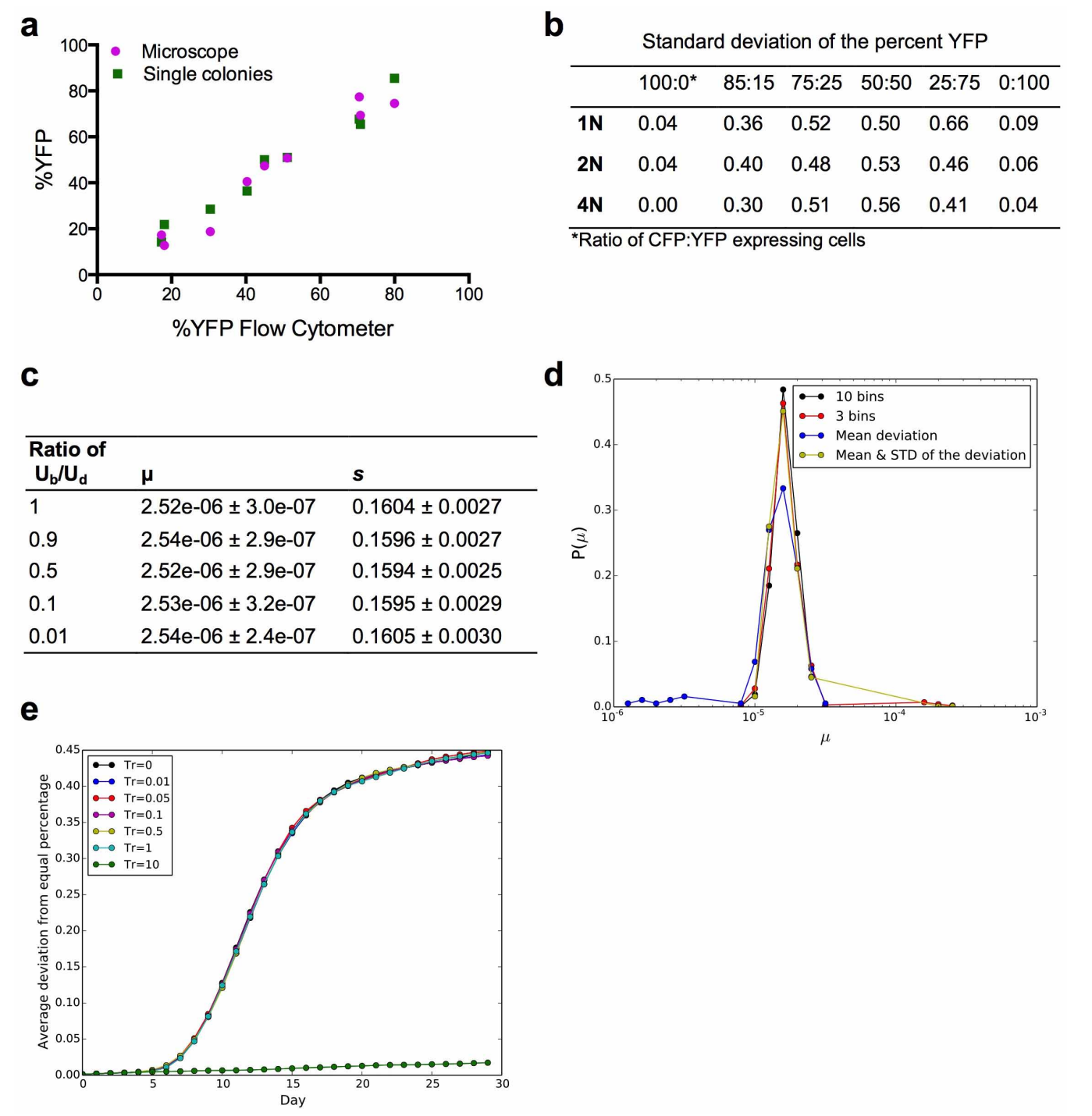
a

Model and Distribution	1N		2N		4N	
	μ	s	μ	s	μ	s
New Exponential	$2.5 \cdot 10^{-3}$ [$3.98 \cdot 10^{-4}$, $3.98 \cdot 10^{-3}$]	0.014 [0.012, 0.018]	$2.5 \cdot 10^{-3}$ [$1 \cdot 10^{-3}$, $3.16 \cdot 10^{-3}$]	0.024 [0.022, 0.026]	$5.0 \cdot 10^{-3}$ [$1.99 \cdot 10^{-3}$, $7.9 \cdot 10^{-3}$]	0.03 [0.028, 0.035]
New Uniform	$2.0 \cdot 10^{-5}$ [$1 \cdot 10^{-5}$, $3.2 \cdot 10^{-5}$]	0.06 [0.115, 0.13]	$7.9 \cdot 10^{-6}$ [$6.3 \cdot 10^{-6}$, $1 \cdot 10^{-5}$]	0.0925 [0.18, 0.195]	$1.2 \cdot 10^{-4}$ [$7.94 \cdot 10^{-5}$, $2.5 \cdot 10^{-4}$]	0.15 [0.275, 0.325]
New Delta	$3.9 \cdot 10^{-6}$ [$2.5 \cdot 10^{-6}$, $6.3 \cdot 10^{-6}$]	0.11 [0.105, 0.115]	$2.5 \cdot 10^{-6}$ [$1.9 \cdot 10^{-6}$, $3.1 \cdot 10^{-6}$]	0.16 [0.155, 0.165]	$1.99 \cdot 10^{-5}$ [$1.58 \cdot 10^{-5}$, $3.1 \cdot 10^{-5}$]	0.31 [0.295, 0.325]
Equivalence Principle model, Delta	$2.5 \cdot 10^{-7}$ [$1.58 \cdot 10^{-7}$, $3.16 \cdot 10^{-7}$]	0.1 [0.95, 0.105]	$2.5 \cdot 10^{-7}$ [$1.99 \cdot 10^{-7}$, $2.81 \cdot 10^{-7}$]	0.13 [0.125, 0.135]	$7.9 \cdot 10^{-7}$ [$6.3 \cdot 10^{-7}$, $1 \cdot 10^{-6}$]	0.2 [0.195, 0.21]

b**c****d****e****f**

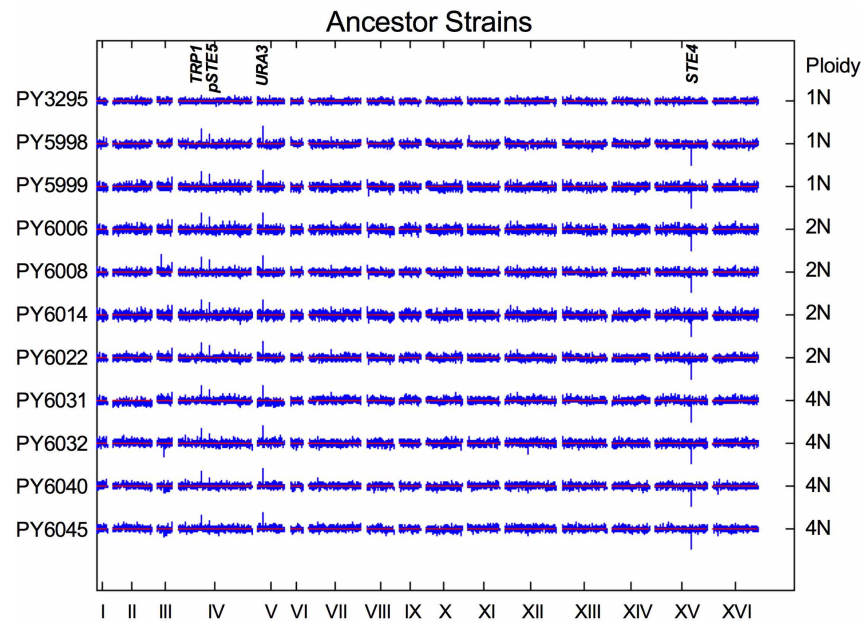
Extended Data Figure 2 | Estimates from our mathematical modelling of the best-fit value of the beneficial mutation rate (μ) and the selection coefficient (s) of each ploidy evolution experiment. **a**, Table of μ and s values that had the best-fit between the simulations and the experimental data; brackets, 95% confidence intervals. Values were determined on the basis of different assumptions about the underlying distribution of beneficial mutations, which included **(b)** uniform, **(c)** delta, and **(d)** exponential distributions. Estimates of μ and s were also obtained with **(e)** the equivalence principle model¹⁷, which assumes a delta distribution of beneficial mutations. Each two-dimensional plot includes the error range obtained by parametric bootstrap of 1,000 independent simulated data sets (Methods). The 95% confidence intervals of μ and s from those 1,000 data sets were defined as the error ranges. **f**, Schematic diagram of the three distributions of fitness effects that we used in our mathematical modelling: exponential (red), uniform (black), and delta (green) distributions. For illustration, we also provide a narrow Gaussian distribution (blue) that is close to a delta function. The real distribution of fitness effects probably has a more complex structure than any of

the examples shown. The diagram illustrates the fact that the shape of the assumed distribution mandates differences in mutation rates. For example, if the mutations that mainly drive adaptation fall within the region of the double arrow, only a small proportion of the mutations from the exponential distribution will fall within this range, necessitating a much higher mutation rate to generate mutations in this region. By contrast, the delta distribution lies in the middle of the double arrow range; therefore, all of the mutations that arise from this distribution are strong enough to contribute to adaptation, resulting in a relatively lower mutation rate. The uniform distribution is intermediate between these two extremes. Only a small portion of the mutations of the uniform distribution is within the double arrow region, but the probability of these mutations is orders of magnitude larger than the exponential. Therefore, the mutation rate of the uniform is closer to the delta than to the exponential distribution. The values used to generate this figure are the best-fit values of μ and s of the haploid populations in the different three distributions. See Methods for more details.



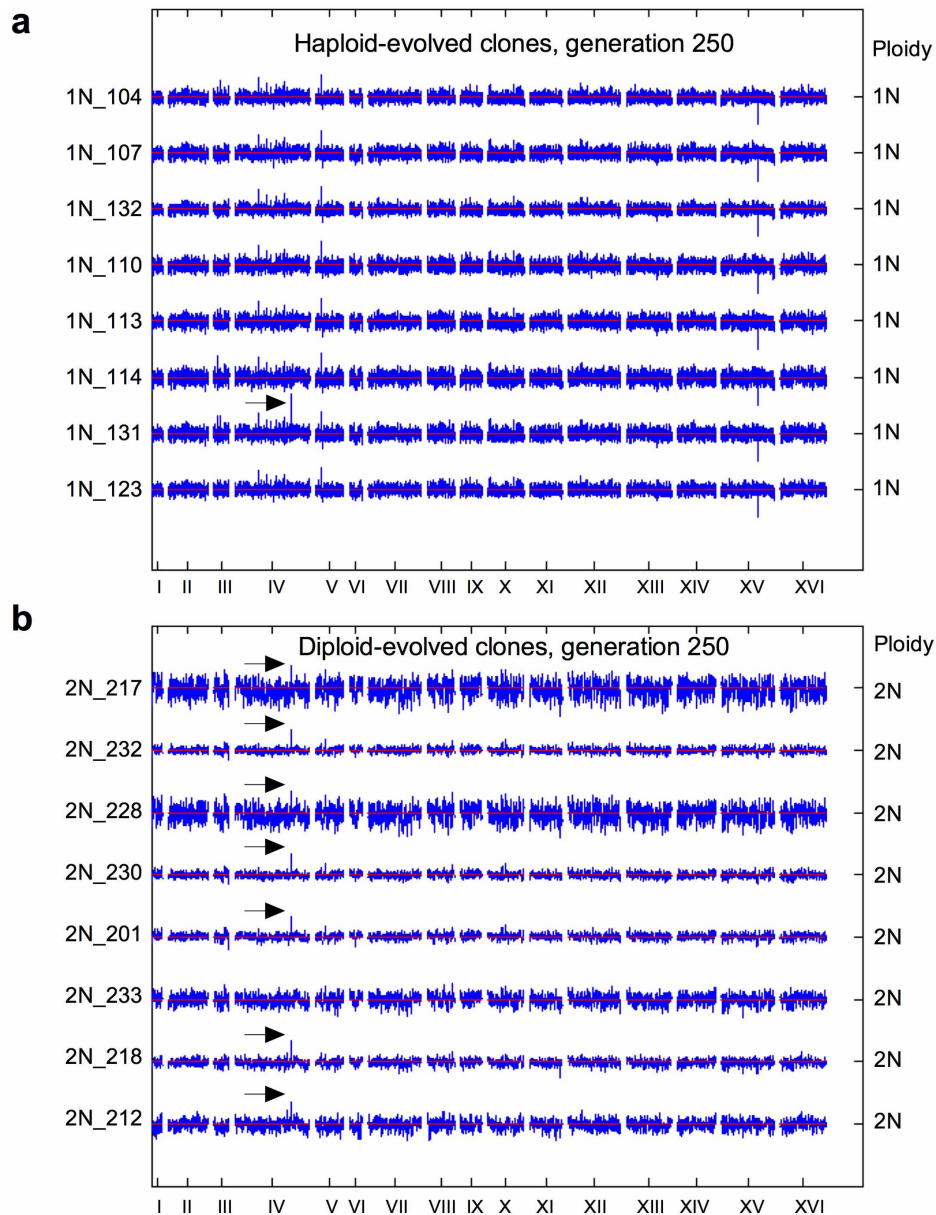
Extended Data Figure 3 | Experimental and computational analyses of the noise in our experimental measurements and of the methods used in our mathematical modelling (see Supplementary Information). **a**, Three different methods were used to determine the percentage of YFP-expressing cells in mixtures of the 1N, 2N, and 4N CFP and YFP ancestor strains. Cells were analysed by flow cytometry (10,000 cells) and fluorescence microscopy (300 cells), and by single colony analysis (96 colonies) of the mixture before galactose induction. The percentage YFP determined by all three methods was highly correlated (Pearson correlation coefficient = 0.98). **b**, Table showing variation in flow cytometry replicate measurements: the standard deviation of the percentage YFP obtained from 48 replicate populations of six different CFP:YFP ratios, for each ploidy type. **c**, Table showing the average and standard deviation of the best-fit values for different ratios between beneficial (U_b) and deleterious (U_d) mutations, obtained from 100 independent data sets. **d**, Evaluation of different summary statistics by calculating the distribution of best-fit values from 1,000 replicate simulations. Four different summary

statistics were used to analyse 1,000 replicates of a parameter pair, s and μ (see Methods). The summary statistic using ten bins has the highest mode and no outliers, and was used to generate our best-fit values. **e**, Criteria for exclusion of near-neutral mutations for implementation of the branching evolutionary model with an exponential distribution of mutations. Shown is the average deviation from equal percentages of YFP and CFP-expressing cells with different thresholds for neutral mutations. The threshold (Tr) represents the fraction of the average fitness effect (s), meaning every mutation whose absolute value was smaller than $Tr \times s$ was excluded. For this scenario (with parameters $\mu = 2 \times 10^{-5}$ and $s = 0.08$), we can exclude every mutation with a fitness effect smaller than s (that is, $Tr = 1$, light blue) without changing the outcome relative to excluding no mutations ($Tr = 0$). However, when excluding all mutations with fitness effects smaller than $10 \times s$ ($Tr = 10$, dark green), the result changes substantially. Thus, for high mutation rates ($N\mu > 1$), we can exclude weak mutations²⁰.



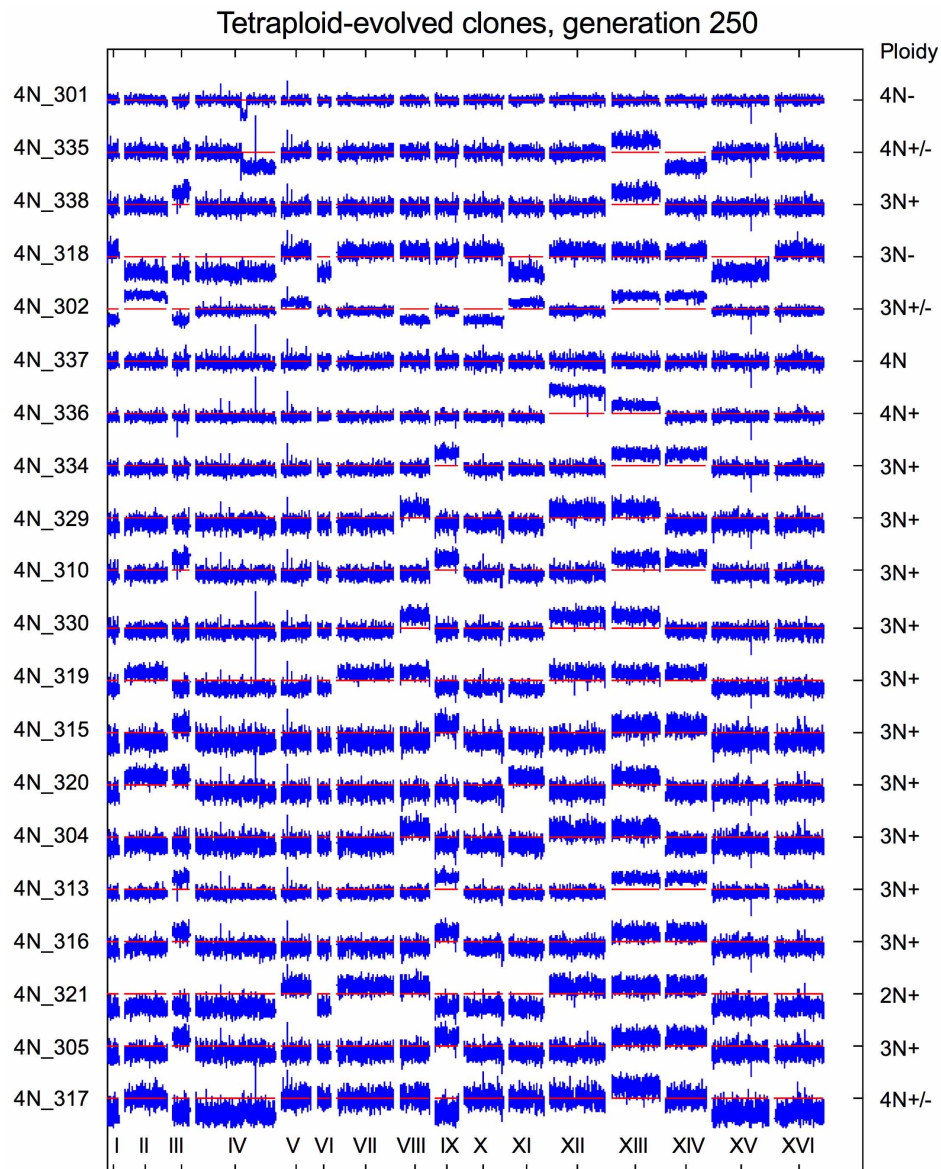
Extended Data Figure 4 | aCGH karyotype of the ancestor strains used in this study. Aneuploidy was not detected in the parental 1N, 2N, or 4N strains. Genomic DNA from each strain was compared with that of an isogenic ancestor PY3295 (BY4741 *MATa ura3 his3 trp1 leu2 LYS2*) and log₂ DNA copy number ratios were plotted using a custom Matlab script. To account for

regions of complete deletion, the data were cropped at log₂ ratios of ± 2.0 and averaged across each chromosome using a sliding window of nine oligonucleotides. A log₂ ratio of zero is indicated by the red line. Loci altered during strain construction are indicated (*TRP1*, *pSTE5*, *URA3*, *STE4*). Strain ploidy, determined by flow cytometry, is indicated on the right.



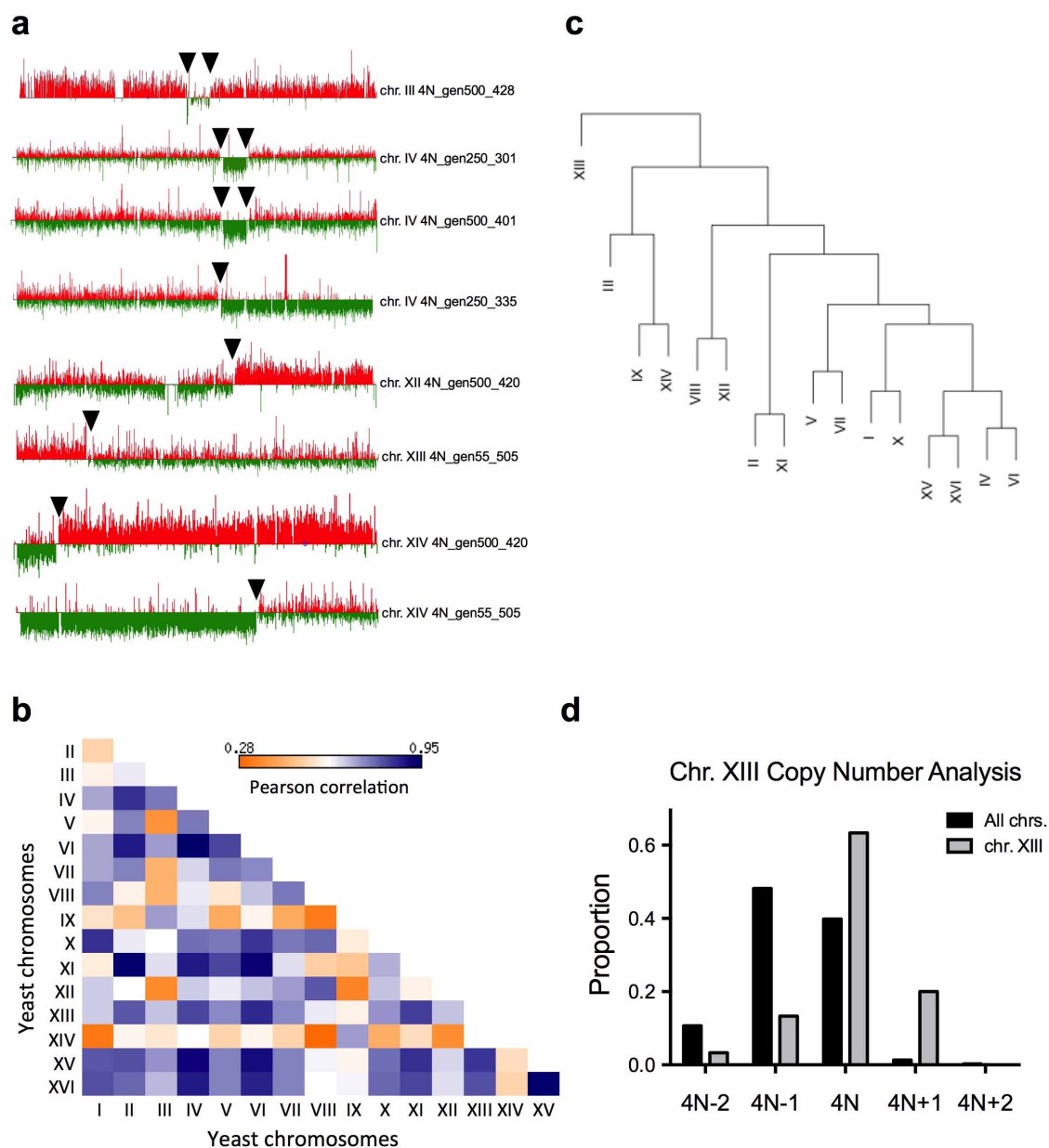
Extended Data Figure 5 | aCGH karyotype of haploid- and diploid-evolved clones at generation 250. **a**, aCGH of eight haploid-evolved clones. Data are displayed as in Extended Data Fig. 4. No aneuploidy was detected. Clone 1N_131 acquired the *HXT6/7* amplification (arrow). **b**, aCGH of eight

diploid-evolved clones. No aneuploidy was detected, but all clones except 2N_233 acquired the *HXT6/7* amplification. Log₂ ratios were averaged across each chromosome using a sliding window of 29 oligonucleotides. The ploidy of the evolved clone, determined by flow cytometry, is indicated on the right.



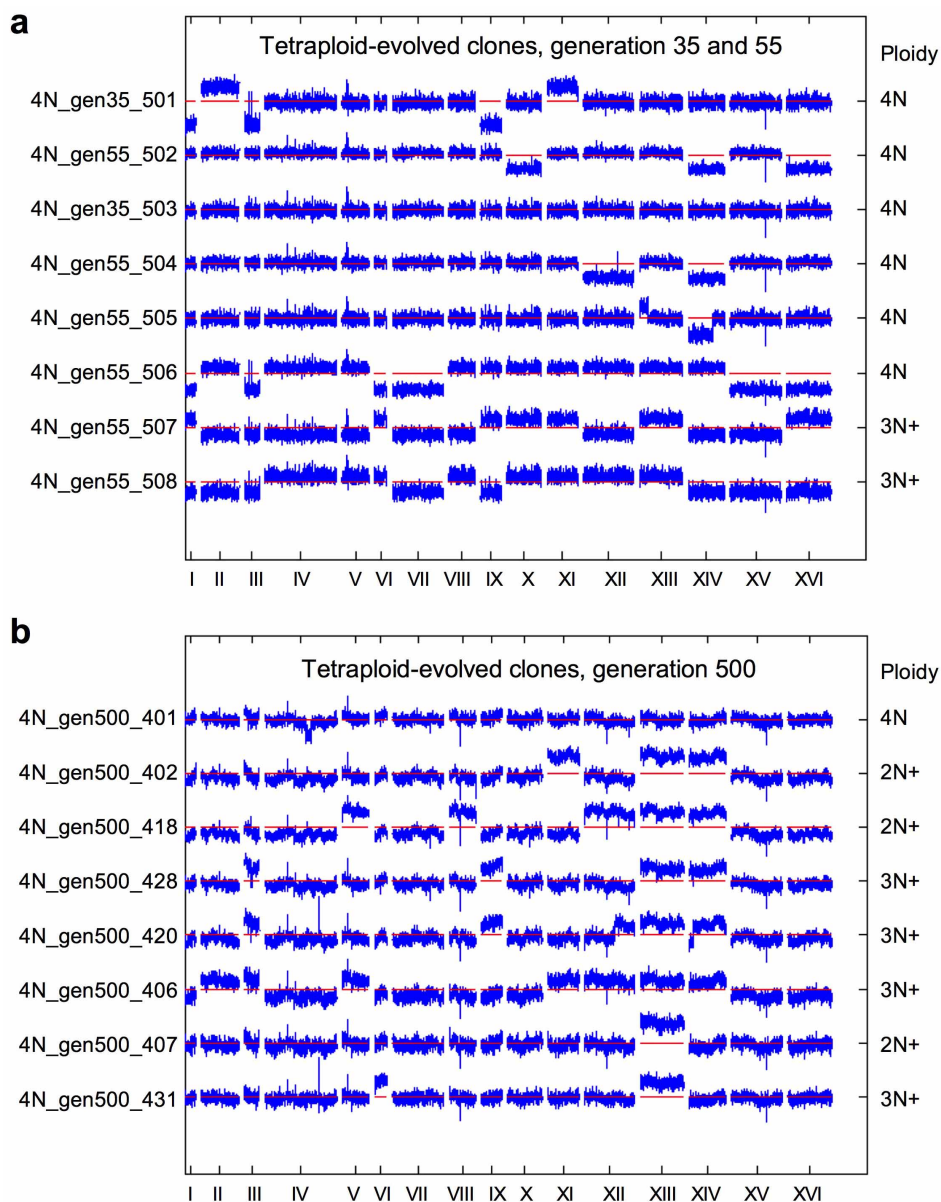
Extended Data Figure 6 | aCGH karyotype for 20 tetraploid-evolved clones at generation 250. aCGH data are displayed as in Extended Data Fig. 4. Note that whole chromosome or large segmental chromosome gain and loss events are observed in all clones except clone 4N_337. Ploidy of the evolved clone,

determined by flow cytometry, is indicated on the right, with +/- indicating chromosome aneuploidy. Some highly aneuploid clones had widely different chromosome copy numbers for different chromosomes (for example, some chromosomes were disomic, others trisomic and tetrasomic).



Extended Data Figure 7 | Analysis of recurrent and concerted chromosome loss events in the tetraploid-evolved clones. **a**, Evolved tetraploids acquired large segmental aneuploidies (regions greater than the ~ 7 kb *HXT6/7* amplification); aCGH data for individual chromosomes with large segmental aneuploidies in 4N-evolved clones (plotted using Treeview⁵²). All breakpoints occurred at or near Ty sequences (arrowheads). **b**, The pairwise patterns (Pearson correlation) of all chromosome copy number alterations in the 4N-evolved clones at generation 250 ($n = 30$, Supplementary Table 2). The copy numbers of some chromosomes were correlated (for example,

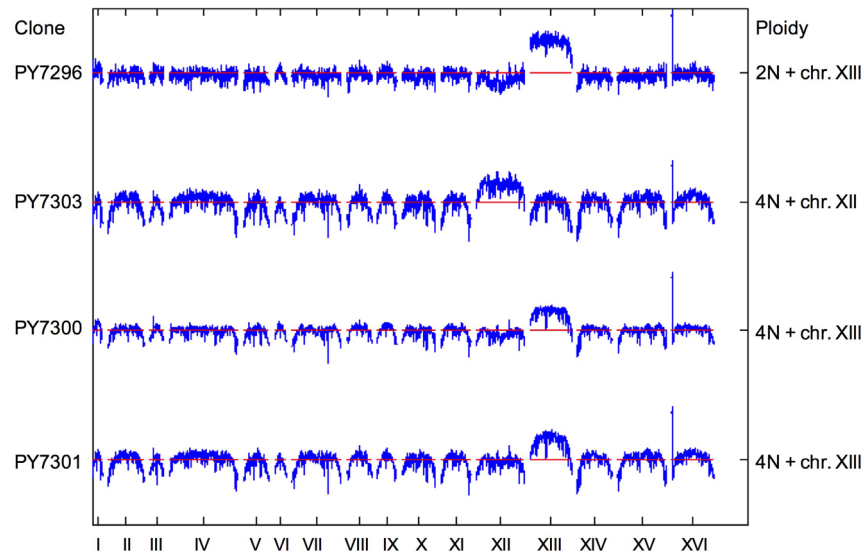
chromosome XV and chromosome XVI), whereas others were anti-correlated (for example, chromosome VIII and chromosome IX), possibly reflecting the need for gene expression balance. **c**, Hierarchical clustering showing the copy number relationship among the chromosomes. **d**, Proportion of all chromosomes in the evolved tetraploid clones with the indicated copy number (black). The copy number of chromosome XIII (grey) in the 4N-evolved clones at generation 250 was significantly different from that of all other aneuploid chromosomes (Cochran–Armitage test, $P < 1 \times 10^{-7}$).



Extended Data Figure 8 | aCGH karyotype for tetraploid-evolved clones at generations 35, 55, and 500. All 4N-evolved clones at (a) generations 35 and 55 and (b) generation 500 were aneuploid for multiple chromosomes or carried large segmental chromosome aneuploidies, except for clone

4N_gen35_503, which remained tetraploid. Data are displayed as in Extended Data Fig. 4. Ploidy of the evolved clone, determined by flow cytometry, is indicated on the right, with +/− indicating chromosome aneuploidy.

Aneuploid Strain Construction



Extended Data Figure 9 | aCGH from isogenic 2N and 4N strains with an extra copy of chromosome XIII or chromosome XII. Data are displayed as in Extended Data Fig. 5b.

Extended Data Table 1 | Yeast strains and plasmids used in this study

Strain (Ploidy) or Plasmid	Parental strain	Relevant genotype	Source
BY3295 (1N)	BY4741	<i>MATa his3Δ leu2Δ met15Δ ura3Δ</i>	Pellman collections
PY5997 (1N)	BY3295	<i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R</i>	This study
PY5998 (1N)	PY5997	<i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY5999 (1N)	PY5997	<i>matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6006 (2N)	PY5999	<i>(2x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6008 (2N)	PY5998	<i>(2x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6014 (2N)	PY5999	<i>(2x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6022 (2N)	PY5998	<i>(2x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6031 (4N)	PY6008	<i>(4x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6032 (4N)	PY6022	<i>(4x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-ceCFP-tADH-SpHIS5</i>	This study
PY6040 (4N)	PY6006	<i>(4x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY6045 (4N)	PY6014	<i>(4x) matΔ::pSTE5-ura3::TRP1 ste4Δ::Hygro^R trp1::Nat^R::pGAL-eYFP-tADH-SpHIS5</i>	This study
PY7284 (1N)	PY5999	<i>SNF3-G439E</i>	This study
PY7289-PY7290 (2N)	PY5999	<i>SNF3-G439E/SNF3</i>	This study
PY7285-PY7288 (2N)	PY5999	<i>SNF3-G439E/SNF3-G439E</i>	This study
PY7291-PY7294 (4N)	PY5999	<i>SNF3-G439E/SNF3/SNF3/SNF3</i>	This study
PY7295 (2N)	S288c	<i>RLY4737 MATa/α ura3Δ his3Δ trp1Δ leu2Δ</i>	Ref. 50
PY7296 (2N)	PY7295	<i>RLY4888 MATa/α + chr. XIII trisomy</i>	Ref. 50
PY7297-PY7299 (4N)	PY7295	<i>MATa/a/α/α</i>	This study
PY7300-PY7302 (4N)	PY7295	<i>MATa/a/α/α + chr. XIII pentasomy</i>	This study
PY7303-PY7305 (4N)	PY7295	<i>MATa/a/α/α + chr. XII pentasomy</i>	This study
PB1500		<i>YFP-tADH-SpHIS5, Amp^R</i>	Yeast Resource Center
PB1499		<i>CFP-tADH Kan^R Amp^R</i>	Yeast Resource Center
PB2452		<i>CFP-tADH SpHIS5, Amp^R</i>	Pellman collection
PB2694		<i>pGAL1-ceCFP-tADH-SpHIS5, Amp^R</i>	This study
PB2697		<i>pGAL1-eYFP-tADH-SpHIS5, Amp^R</i>	This study
PB2314		<i>MATa::pSTE5-URA3, Amp^R</i>	Ref. 12
PB1308		<i>ura3::TRP1 Amp^R</i>	Ref. 46
PB2577		<i>MATa URA3 CEN Amp^R</i>	This study
B1819		<i>LEU2 CEN Amp^R</i>	Pellman collection
PB2647		<i>STE4 LEU2 CEN Amp^R</i>	This study
PB2649		<i>STE4 URA3 CEN MATα, Amp^R</i>	This study
PB1640		<i>hphMX4 Amp^R</i>	Ref. 47
PB1942		<i>pGAL-HO HIS3 Amp^R</i>	Gift of the Fink lab
PB1650		<i>pGAL-HO URA3 LEU2</i>	Gift of the Elion lab
pCORE		<i>kanMX4 KIURA3</i>	Ref. 48

Thermosensory processing in the *Drosophila* brain

Wendy W. Liu¹, Ofer Mazar^{1,2} & Rachel I. Wilson¹

In *Drosophila*, just as in vertebrates, changes in external temperature are encoded by bidirectional opponent thermoreceptor cells: some cells are excited by warming and inhibited by cooling, whereas others are excited by cooling and inhibited by warming^{1,2}. The central circuits that process these signals are not understood. In *Drosophila*, a specific brain region receives input from thermoreceptor cells^{2,3}. Here we show that distinct genetically identified projection neurons (PNs) in this brain region are excited by cooling, warming, or both. The PNs excited by cooling receive mainly feed-forward excitation from cool thermoreceptors. In contrast, the PNs excited by warming ('warm-PNs') receive both excitation from warm thermoreceptors and crossover inhibition from cool thermoreceptors through inhibitory interneurons. Notably, this crossover inhibition elicits warming-evoked excitation, because warming suppresses tonic activity in cool thermoreceptors. This in turn disinhibits warm-PNs and sums with feed-forward excitation evoked by warming. Crossover inhibition could cancel non-thermal activity (noise) that is positively correlated among warm and cool thermoreceptor cells, while reinforcing thermal activity which is anti-correlated. Our results show how central circuits can combine signals from bidirectional opponent neurons to construct sensitive and robust neural codes.

The simplest model of thermosensory processing is that warm thermoreceptor cells contribute only to the representation of warming, whereas cool cells contribute only to the representation of cooling (the labelled-line hypothesis^{4,5}). In support of this idea, ablating heat-activated thermoreceptor cells in the mouse abolishes behavioural responses to noxious warming but not cooling⁶. Similarly, in *Drosophila*, silencing warm-activated thermoreceptor cells attenuates behavioural responses to warming but not cooling; conversely, silencing cool-activated thermoreceptor cells attenuates behavioural responses to cooling but not warming². An alternative model is that different thermoreceptor cells interact centrally, so that a single cell contributes to the representation of both cooling and warming^{4,5}. This idea could explain perceptual phenomena like the thermal grill illusion, where interlaced warm and cool objects produce the perception of burning heat⁷.

We investigated these possibilities in the *Drosophila* thermosensory system. Peripheral thermoreceptor cells in *Drosophila* reside mainly in the antenna^{8–10}, and are housed inside two structures termed the arista and the sacculus^{2,3}. The arista contains three cells that are excited by warming and inhibited by cooling ('warm cells'), as well as three cells that are excited by cooling and inhibited by warming ('cool cells'). The sacculus contains additional cool cells. All these cells project to a region of the brain called the proximal antennal protocerebrum¹¹. Recent studies have identified several genes involved in thermal transduction in the antenna^{2,3,10,12}. Among these, the warm receptor *Gr28b.d* has a particularly specific expression pattern^{3,13}, permitting genetic access to the warm cells in the arista.

In a visual screen of ~7,000 Gal4 lines¹⁴, we identified several that label central neurons with dendrites in the proximal antennal protocerebrum and axons projecting to higher brain regions. We call such cells thermosensory projection neurons (PNs). We used these Gal4 lines to drive GFP expression in small groups of PNs and performed *in vivo* whole-cell current-clamp recordings from these neurons. We included

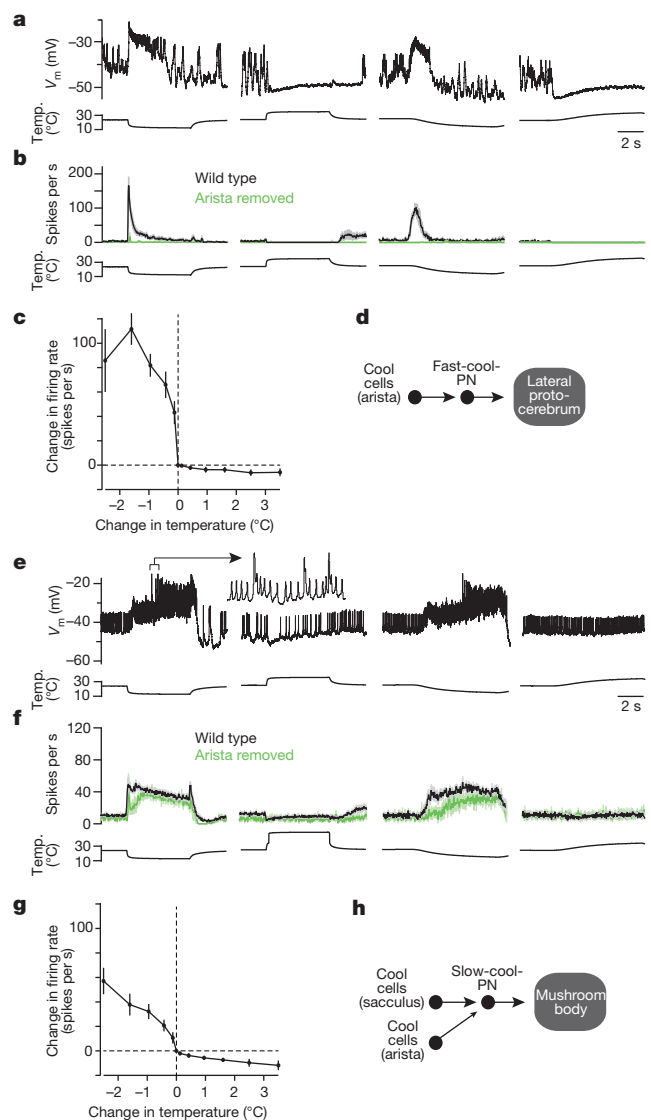


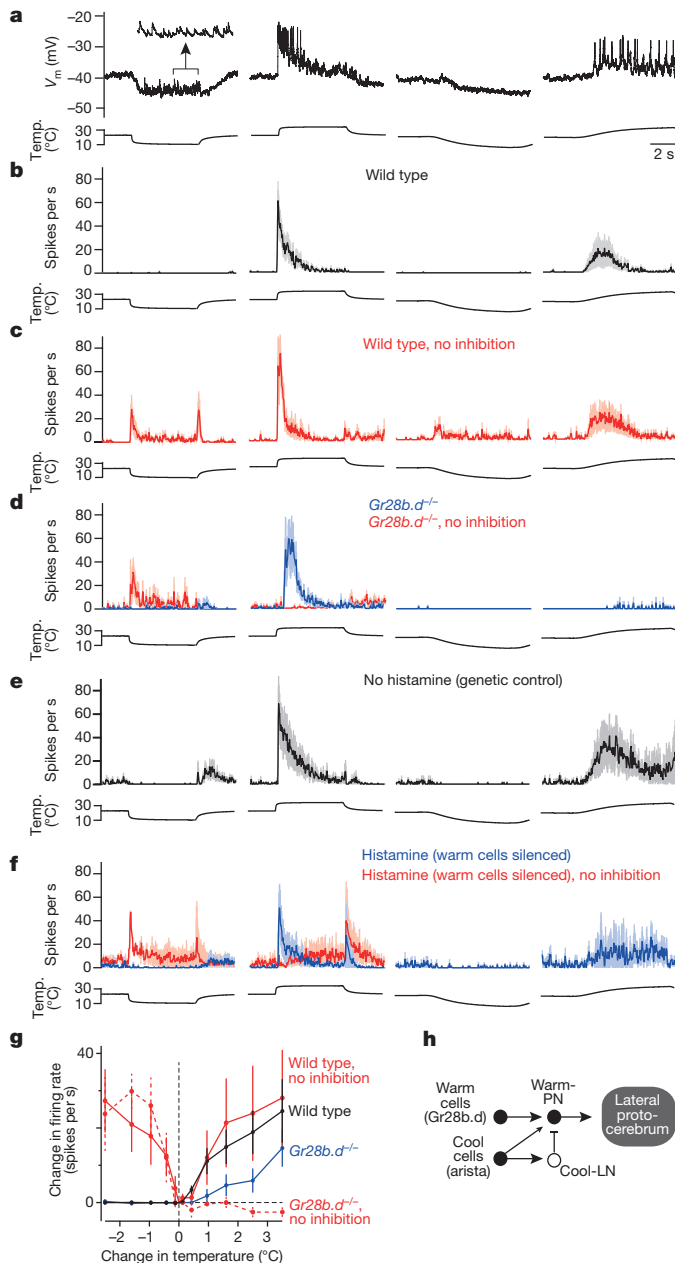
Figure 1 | Projection neurons excited by cooling. **a–d**, Fast-cool-PNs. **a**, *In vivo* recording from a fast-cool-PN. Stimuli are steps from ~24 °C to ~13 °C or ~34 °C. V_m , membrane potential. **b**, Mean firing rate for the fast-cool-PNs \pm s.e.m. across experiments ($n = 7–10$ PNs per stimulus per condition), with either arista intact or arista removed. **c**, Sensitivity to small temperature steps: mean change in firing rate versus change in temperature, mean \pm s.e.m. ($n = 5$). **d**, Proposed circuit for fast-cool-PNs (see Supplementary Discussion regarding circuit diagrams). **e–h**, Slow-cool-PNs. **e**, Slow-cool-PNs show less adaptation, and they respond with similar peak firing rates to fast and slow cooling. Prolonged depolarization can elicit complex spikes (inset, expanded 12-fold). **f**, For the slow-cool-PNs, removing the arista has comparatively little effect ($n = 7–14$); **g**, Slow-cool-PNs are less sensitive to small cooling steps ($n = 7$). In **h** the thinner arrow represents a weaker connection.

¹Department of Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA. ²Harvard NeuroDiscovery Center, Harvard Medical School, 220 Longwood Avenue, Boston, Massachusetts 02115, USA.

biocytin in the patch pipette to reconstruct the morphology of individual neurons.

Because peripheral thermoreceptor cells reside in the antenna, we generated a thermal stimulus by manipulating the temperature of a stream of solution flowing over the antenna while keeping the temperature of the brain constant (Extended Data Fig. 1). As a control, we confirmed that cutting the axons of peripheral thermoreceptor cells abolished responses in this brain region (see Methods). We used large temperature changes to robustly activate peripheral thermoreceptor cells ($\sim 10^\circ\text{C}$). These large temperature changes were delivered at two speeds: fast ($\sim 100\text{ ms}$) and slow ($\sim 7\text{ s}$). We also used small steps ($<1^\circ\text{C}$ over $\sim 100\text{ ms}$) to measure sensitivity near threshold. This panel of stimuli revealed four physiological types of PNs, each with a distinctive morphology (Extended Data Fig. 2).

Two PN types were excited by cooling and inhibited by warming. One type showed strong adaptation to sustained temperature decreases (Fig. 1a, b and Extended Data Fig. 3). These PNs also responded better to fast cooling than to slow cooling. We term these ‘fast-cool-PNs’. Fast-cool-PNs were highly sensitive to small cooling steps (Fig. 1c).



Fast-cool-PNs project to the lateral protocerebrum, a higher brain region that receives multimodal inputs¹⁵.

By contrast, the second type of ‘cool-PN’ showed little adaptation to large sustained temperature decreases (Fig. 1e, f and Extended Data Fig. 3). Moreover, slow and fast cooling elicited similar peak firing rates (Fig. 1e, f). We term these ‘slow-cool-PNs’. Slow-cool-PNs project to the mushroom body, a higher brain region involved in learning and memory¹⁶. Slow-cool-PNs were less sensitive than fast-cool-PNs to small cooling steps (Fig. 1c, g). Thus, these two cool-PN types encode different features of cool stimuli (Supplementary Discussion).

These cool-PN types have dendrites that overlap with the axons of the cool thermoreceptor cells (Extended Data Fig. 4). Clipping the arista largely eliminated the responses of fast-cool-PNs (Fig. 1b). By contrast, the responses of slow-cool-PNs remained mostly intact, although the initial response to cooling was diminished, and tonic firing at room temperature was reduced (Fig. 1f). All responses in slow-cool-PNs were abolished by cutting the antennal nerve (data not shown). Together, these results suggest a simple wiring scheme: fast-cool-PNs receive input from the cool cells in the arista, whereas slow-cool-PNs receive input mainly from the cool cells in the sacculus, with a weaker input from cool cells in the arista.

The idea of a simple wiring scheme was reinforced by additional experiments on the cool-PNs. Namely, we observed only slight disinhibition in the cooling responses of cool-PNs when we bath-applied antagonists of inhibitory neurotransmitter receptors (Extended Data Fig. 5). Warming still suppressed the tonic activity of the cool-PNs when synaptic inhibition was blocked, implying that warming can inhibit these neurons by suppressing tonic excitation from cool thermoreceptor cells. Cool-PN responses were essentially unaltered by a functional null mutation in *Gr28b.d*, the warm receptor in the arista³ (Extended Data Fig. 5). Together, these results suggest that the thermal responses of the cool-PNs are driven primarily by cool thermoreceptor cells, with little role for warm thermoreceptor cells (Fig. 1d, h).

Unlike the representation of cooling, we found that the representation of warming is mediated by a relatively complex circuit. We identified one PN type that is excited by warming and inhibited by cooling (Fig. 2a, b). These warm-PNs project to the lateral protocerebrum (Extended Data Fig. 2). Warm-PNs responded better to fast warming than to slow warming, and they showed adaptation to prolonged temperature increases (Fig. 2b and Extended Data Fig. 3). Their responses were abolished by removing the arista (data not shown).

Blocking synaptic inhibition depolarized warm-PNs (by $6 \pm 1\text{ mV}$, mean \pm s.e.m.), suggesting these cells are subject to tonic inhibition.

Figure 2 | Projection neurons excited specifically by warming. **a**, A warm-PN. Cooling evokes a barrage of excitatory postsynaptic potentials (inset, expanded fourfold). V_m , membrane potential. **b**, Mean firing rate for warm-PNs ($n = 6-8$). **c**, Mean firing rate with inhibition blocked using picrotoxin (an antagonist of GABA_A and GluCl receptors^{20,25}) and CGP54626 (a GABA_B antagonist²⁵) ($n = 3$ or 4). Spiking at cooling onset suggests excitation from the cool pathway. Spiking at cooling offset likely arises from the warm pathway and is normally masked by slowly-decaying cooling-evoked inhibition. **d**, Mean firing rate in the *Gr28b.d* mutant ($n = 5-7$). The response to slow warming was smaller than that in wild type, although the difference was not statistically significant ($P = 0.20$). Blocking inhibition eliminates the excitatory response to fast warming steps ($n = 4$). **e**, Mean firing rate in genetic background for histamine experiments, before injecting histamine ($n = 7$). A histamine-gated chloride channel is expressed in warm thermoreceptor cells. **f**, Injecting histamine to silence warm thermoreceptor cells significantly reduces responses to slow warming ($n = 7$, $P < 0.05$, paired t -test). Blocking inhibition eliminates warming-evoked excitation and unmasks cooling-evoked excitation ($n = 6$). See Extended Data Fig. 7 for controls. **g**, Sensitivity to small temperature steps (wild type $n = 9$; wild type with inhibition blocked $n = 6$; *Gr28b.d* mutant $n = 5$, *Gr28b.d* mutant with inhibition blocked $n = 3$). In wild type, blocking inhibition had no significant effect on sensitivity to warming (unpaired t -tests, with iterative Bonferroni corrections; see Methods). However, responses were significantly smaller in the mutant near the threshold of warming-evoked excitation ($+0.425^\circ\text{C}$, $P = 0.017$). **h**, Proposed circuit for the warm-PNs.

Notably, blocking synaptic inhibition also unmasked weak cooling-evoked excitation in warm-PNs (Fig. 2c). This implies that these PN receive input from both warm and cool thermoreceptor cells in the arista, and weak cooling-evoked excitation is normally masked by strong cooling-evoked inhibition. Moreover, although cooling hyperpolarizes these cells, it also elicits a barrage of excitatory postsynaptic potentials riding on the hyperpolarized membrane potential (Fig. 2a, inset). This observation is further evidence that these cells receive excitatory input from cool thermoreceptor cells. Finally, we note that the dendrites of warm-PNs overlap with the axons of both warm and cool thermoreceptor cells (Extended Data Fig. 4), compatible with the idea that they receive input from both.

Notably, mutating the warm receptor gene *Gr28b.d* did not abolish excitatory responses to warming in these PN (Fig. 2d). Indeed, responses to fast warming were almost unchanged. Interestingly, however, responses to slow warming were smaller than in wild-type PN. These results suggest that there are two pathways for generating warming-evoked excitation in these PN. *Gr28b.d* is necessary for generating normal responses to slow warming, but fast warming recruits a second pathway independent of *Gr28b.d* (Supplementary Discussion).

Based on these results, we hypothesized that the cool pathway can excite these neurons in response to warming through disinhibition. This hypothesis is motivated by the observation that cool thermoreceptor cells are tonically active and their tonic activity is suppressed by warming². An interposed inhibitory neuron could relay this tonic activity in the form of tonic inhibition, and warming would withdraw this inhibition. This hypothesis predicts that blocking synaptic inhibition will eliminate the second pathway. Indeed, we found that blocking inhibition abolished warming-evoked excitation in warm-PNs in the *Gr28b.d* mutant (Fig. 2d). This result was not observed in a wild-type background; in that case, blocking inhibition potentiated warming responses rather than eliminating them (Fig. 2c). As an additional control, we confirmed that blocking inhibition had essentially no effect on the cool-PNs in the *Gr28b.d* mutant (Extended Data Fig. 5).

In principle, a mutation in the warm receptor *Gr28b.d* might cause the circuit to re-wire, or there might be other receptor molecules that could mediate responses in warm thermoreceptor cells when *Gr28b.d* is absent. To address these caveats, we acutely silenced the warm thermoreceptor cells themselves. We used the LexA system to selectively express a histamine-gated chloride channel (Ort) in warm thermoreceptor cells under the control of the *Gr28b.d* promoter (Extended Data Fig. 6). Neurons that mis-express Ort in this manner can be acutely silenced by histamine¹⁷. We first confirmed that simply expressing Ort in warm thermoreceptor cells did not substantially alter the responses of warm-PNs (Fig. 2e). During these recordings, we then microiontophoresed histamine into the antenna to silence the warm thermoreceptor cells. PN responses to fast warming were largely intact, although responses to slow warming were significantly reduced (Fig. 2f). Blocking inhibition nearly abolished warming-evoked excitation when the warm thermoreceptor cells were silenced (Fig. 2f). Thus, this manipulation phenocopies the *Gr28b.d* mutant. These results further support a model whereby the cool pathway can cause warming-induced excitation via disinhibition.

Small temperature steps ($<1^{\circ}\text{C}$) yielded further insight into the two pathways which provide input to warm-PNs. Under normal conditions, warm-PNs were excited by small warming steps (Fig. 2g). Blocking inhibition unmasked a response to small cooling steps (Fig. 2g). Responses to small warming steps were still present in the *Gr28b.d* mutant, but—unlike responses to large warming steps—were significantly reduced. Responses to small warming steps were eliminated in the mutant after inhibition was blocked (Fig. 2g). These results bolster the conclusion that both the warm pathway and the cool pathway can drive excitatory responses in warm-PNs (Fig. 2h). The warm pathway is most important when the rate of warming is low ($\sim 1^{\circ}\text{C}$ per s, as in our ‘small’ stimuli and ‘slow large’ stimuli). The cool pathway becomes dominant when the rate of warming is high ($\sim 10^{\circ}\text{C}$ over 100 ms).

The fourth and final projection neuron type we identified was excited by both cooling and warming (Fig. 3a, b). These ‘warm-cool-PNs’ project to the lateral protocerebrum (Extended Data Fig. 2). They responded to fast but not slow changes in temperature, and their responses to sustained stimuli were strongly adapting (Fig. 3b and Extended Data Fig. 3). Thus, these neurons signal rapid temperature shifts, regardless of the direction of those shifts. Removing the arista eliminated all responses in these neurons (data not shown). Blocking inhibition depolarized these cells (by 8 ± 4 mV, mean \pm s.e.m.) with modest effects on their thermal responses (Fig. 3c).

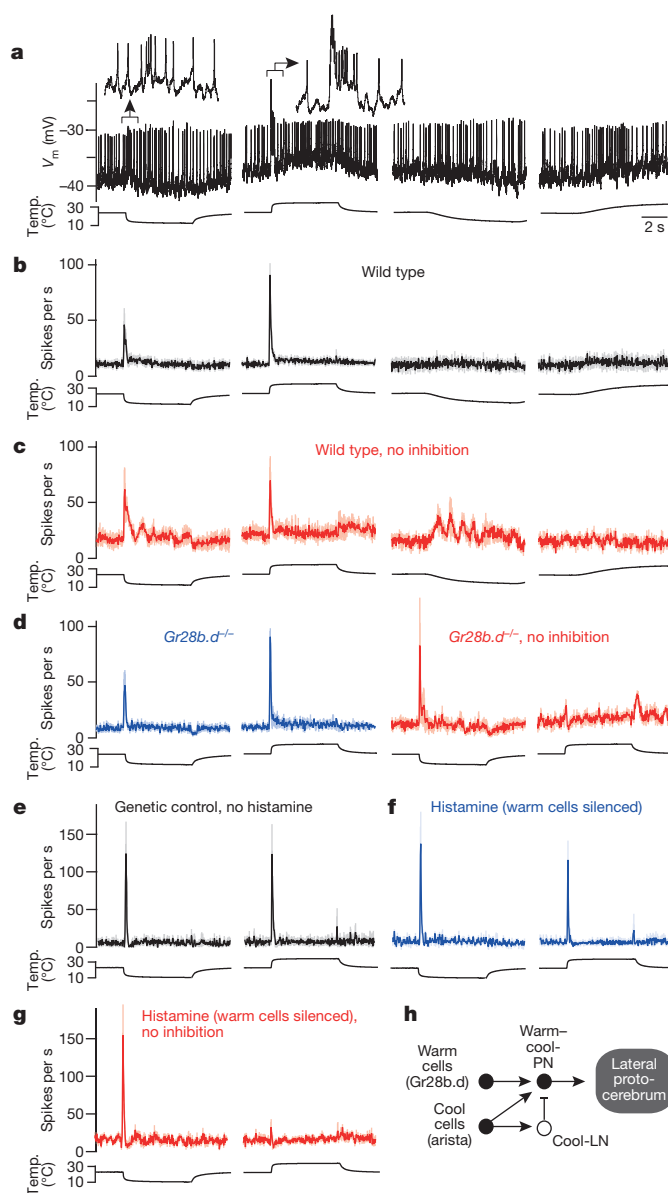


Figure 3 | Projection neurons excited by both warming and cooling.

a, A warm-cool-PN. Fast warming and cooling elicit transient bursts of spikes (insets, expanded sevenfold). Stimulus offset does not elicit a response, probably because temperature is changing more slowly. V_m , membrane potential. **b**, Mean firing rate ($n = 10$ – 13). **c**, Mean firing rate with inhibition blocked ($n = 3$ – 5). **d**, Mean firing rate in the *Gr28b.d* mutant, before and after blocking inhibition ($n = 7$ – 11 and 4). **e**, Mean firing rate in genetic background for histamine experiments, before injecting histamine ($n = 4$). A histamine-gated chloride channel is expressed in warm thermoreceptor cells. **f**, **g**, Responses after injecting histamine to silence warm thermoreceptor cells ($n = 4$) and after blocking inhibition ($n = 3$). See Extended Data Fig. 7 for controls. **h**, Proposed circuit for the warm-cool-PNs.

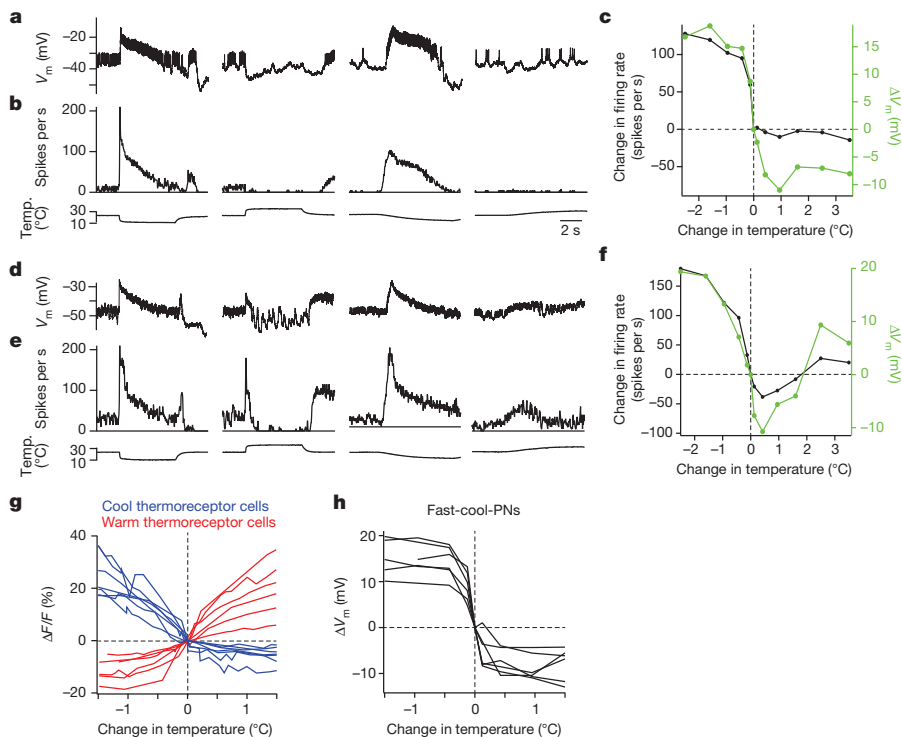


Figure 4 | Inhibitory local neurons responding to thermal stimuli. **a**, Recording from a GABAergic LN that is excited by cooling and inhibited by warming (a 'cool-LN'). V_m , membrane potential. **b**, Trial-averaged firing rates. **c**, Sensitivity to small temperature steps. Neurotransmitter release from these neurons may be related to both membrane potential and firing rate. **d–f**, Same as above, but for a different GABAergic LN. **g**, Sensitivity to small temperature steps in thermoreceptor cells of the arista, measured as the change in GCaMP3 fluorescence in experiments where either warm or cool thermoreceptor cells expressed the calcium indicator. Each line represents a different experiment. See also Extended Data Fig. 6. **h**, Sensitivity to small temperature steps, measured from the membrane potential of fast-cool-PNs, the more sensitive of the two cool-PN types. Each line represents a different experiment. The population of cool thermoreceptors must be at least this sensitive to warming.

Notably, warming-evoked excitation persisted in warm-cool-PNs in the *Gr28b.d* mutant (Fig. 3d and Extended Data Fig. 8), just as we observed in warm-PNs. As before, blocking synaptic inhibition abolished warming-evoked excitation in the *Gr28b.d* mutant (Fig. 3d). Silencing thermoreceptor warm cells with histamine again phenocopied the *Gr28b.d* mutation (Fig. 3e–g). These results imply that warming-evoked excitation in these cells arises from two sources: increased excitation from the warm pathway, and withdrawal of tonic inhibition from the cool pathway (Fig. 3h). The dendrites of warm-cool-PNs overlap with the axons of both warm and cool receptor cells (Extended Data Fig. 4), compatible with the conclusion that they receive input from both.

These findings predict that there are inhibitory neurons in this brain region that are tonically active, inhibited by warming, and excited by cooling. To test these predictions, we made recordings from inhibitory local neurons in this brain region. The main inhibitory neurotransmitters in the *Drosophila* brain are GABA and glutamate^{18–20}, and so we labelled either GABAergic or glutamatergic neurons with GFP to target our electrodes to these cells (Extended Data Fig. 9). In total, we recorded from 17 inhibitory neurons that arbourized locally in this brain region and responded to thermal stimuli (Extended Data Fig. 2).

Most of the inhibitory local neurons (LNs) we recorded from (14 of 17) were tonically active at room temperature, inhibited by warming, and excited by cooling (Fig. 4a–f). We term these 'cool-LNs'. (The three remaining LNs had the opposite tuning profile; Extended Data Fig. 10.) Cool-LNs were found within both the GABAergic and glutamatergic populations.

Several features of PN responses are explained by a model where cool-LNs inhibit the warm-PNs and warm-cool-PNs. First, cooling excites these LNs, which could mask cooling-evoked excitation onto warm-PNs (and also attenuate cooling-evoked excitation in warm-cool-PNs), except when synaptic inhibition is blocked. Conversely, warming inhibits cool-LNs, which should disinhibit PN, allowing them to become excited even when the warm receptor is mutated or warm thermoreceptor cells are silenced, but in a manner that depends on intact synaptic inhibition.

In summary, our results indicate an asymmetry in the central processing of warm versus cool stimuli. Cool-PNs are driven primarily by excitation from cool thermoreceptor cells. These simple circuits evoke the

classic concept of the labelled line. By contrast, warm-PNs and warm-cool-PNs receive excitation from both the warm and cool pathway, as well as potent crossover inhibition from the cool pathway. This more complex arrangement is reminiscent of some circuits in the mammalian spinal cord (Supplementary Discussion).

In warm-PNs (and warm-cool-PNs), crossover inhibition from the cool pathway may help to cancel non-thermal activity (noise) that is correlated among warm and cool thermoreceptor cells. Correlated activity could arise from mechanical forces that bend the arista²¹, thereby deforming and potentially depolarizing both warm and cool cells²². Unlike the thermal responses of these cells, which are anti-correlated, correlated activity in warm and cool thermoreceptor cells would be cancelled by crossover inhibition. Crossover inhibition might also increase sensitivity to small warming steps, because cool thermoreceptor cells are sensitive to warming, even though they are inhibited rather than excited by warming (Fig. 4g, h). Notably, some insects lack warm receptor cells entirely and yet possess both cool-PNs and warm-PNs. It has been suggested that disinhibition might be the only source of warming-evoked excitation in the warm-PNs of these insects^{23,24}. In mammals, many cool peripheral thermoreceptor cells are inhibited by warming, and warm peripheral thermoreceptor cells are inhibited by cooling¹. Thus, the mammalian spinal cord might also use crossover inhibition, because each receptor type is informative about both warming and cooling.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 July; accepted 18 December 2014.

Published online 4 March 2015.

- Schepers, R. J. & Ringkamp, M. Thermoreceptors and thermosensitive afferents. *Neurosci. Biobehav. Rev.* **34**, 177–184 (2010).
- Gallio, M., Ofstad, T. A., Macpherson, L. J., Wang, J. W. & Zuker, C. S. The coding of temperature in the *Drosophila* brain. *Cell* **144**, 614–624 (2011).
- Ni, L. *et al.* A gustatory receptor paralogue controls rapid warmth avoidance in *Drosophila*. *Nature* **500**, 580–584 (2013).
- Ma, Q. Labeled lines meet and talk: population coding of somatic sensations. *J. Clin. Invest.* **120**, 3773–3778 (2010).
- Prescott, S. A. & Ratte, S. Pain processing by spinal microcircuits: afferent combinatorics. *Curr. Opin. Neurobiol.* **22**, 631–639 (2012).

6. Cavanaugh, D. J. *et al.* Distinct subsets of unmyelinated primary sensory fibers mediate behavioral responses to noxious thermal and mechanical stimuli. *Proc. Natl Acad. Sci. USA* **106**, 9075–9080 (2009).
7. Green, B. G. Temperature perception and nociception. *J. Neurobiol.* **61**, 13–29 (2004).
8. Sayeed, O. & Benzer, S. Behavioral genetics of thermosensation and hygrosensation in *Drosophila*. *Proc. Natl Acad. Sci. USA* **93**, 6079–6084 (1996).
9. Zars, T. Two thermosensors in *Drosophila* have different behavioral functions. *J. Comp. Physiol. A* **187**, 235–242 (2001).
10. Tang, X., Platt, M. D., Lagnese, C. M., Leslie, J. R. & Hamada, F. N. Temperature integration at the AC thermosensory neurons in *Drosophila*. *J. Neurosci.* **33**, 894–901 (2013).
11. Chiang, A. S. *et al.* Three-dimensional reconstruction of brain-wide wiring networks in *Drosophila* at single-cell resolution. *Curr. Biol.* **21**, 1–11 (2011).
12. Hamada, F. N. *et al.* An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* **454**, 217–220 (2008).
13. Thorne, N. & Amrein, H. Atypical expression of *Drosophila* gustatory receptor genes in sensory and central neurons. *J. Comp. Neurol.* **506**, 548–568 (2008).
14. Pfeiffer, B. D. *et al.* Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proc. Natl Acad. Sci. USA* **105**, 9715–9720 (2008).
15. Ito, K. *et al.* A systematic nomenclature for the insect brain. *Neuron* **81**, 755–765 (2014).
16. Heisenberg, M. Mushroom body memoir: from maps to models. *Nature Rev. Neurosci.* **4**, 266–275 (2003).
17. Liu, W. W. & Wilson, R. I. Transient and specific inactivation of *Drosophila* neurons *in vivo* using a native ligand-gated ion channel. *Curr. Biol.* **23**, 1202–1208 (2013).
18. Enell, L., Hamasaka, Y., Kolodziejczyk, A. & Nassel, D. R. γ -Aminobutyric acid (GABA) signaling components in *Drosophila*: immunocytochemical localization of GABA_B receptors in relation to the GABA_A receptor subunit RDL and a vesicular GABA transporter. *J. Comp. Neurol.* **505**, 18–31 (2007).
19. Daniels, R. W., Gelfand, M. V., Collins, C. A. & DiAntonio, A. Visualizing glutamatergic cell bodies and synapses in *Drosophila* larval and adult CNS. *J. Comp. Neurol.* **508**, 131–152 (2008).
20. Liu, W. W. & Wilson, R. I. Glutamate is an inhibitory neurotransmitter in the *Drosophila* olfactory system. *Proc. Natl Acad. Sci. USA* **110**, 10294–10299 (2013).
21. Foelix, R. F., Stocker, R. F. & Steinbrecht, R. A. Fine structure of a sensory organ in the arista of *Drosophila melanogaster* and some other dipterans. *Cell Tissue Res.* **258**, 277–287 (1989).
22. Sachs, F. & Morris, C. E. Mechanosensitive ion channels in non-specialized cells. In *Reviews of Physiology, Biochemistry and Pharmacology* Vol. 132 (Springer, 1998).
23. Fischer, H. & Tichy, H. Cold-receptor cells supply both cold- and warm-responsive projection neurons in the antennal lobe of the cockroach. *J. Comp. Physiol. A* **188**, 643–648 (2002).
24. Ruchty, M., Helmchen, F., Wehner, R. & Kleineidam, C. J. Representation of thermal information in the antennal lobe of leaf-cutting ants. *Front. Behav. Neurosci.* **4**, 174 (2010).
25. Wilson, R. I. & Laurent, G. Role of GABAergic inhibition in shaping odor-evoked spatiotemporal patterns in the *Drosophila* antennal lobe. *J. Neurosci.* **25**, 9069–9079 (2005).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to C. -H. Lee for the gift of *LexAOp-HA-Ort* flies, P. Garrity for *Gr28b.d-Gal4*, and A. Rajan for *Gad1-Gal4* (line 2B). A. DiAntonio kindly supplied the anti-dVGluT antibody. We thank the Wilson laboratory, A. Samuel, P. Garrity, L. Griffith and D. Ginty for comments on the manuscript. D. Rogulja's laboratory and M. Y. Wong shared reagents and expertise. A portion of this work was supported by NIH grant R01 DC008174. W.W.L. is supported by an HHMI International Research Fellowship and a Presidential Scholarship from the MD–PhD Program at Harvard Medical School. R.I.W. is a Howard Hughes Investigator.

Author Contributions W.W.L. and R.I.W. conceived the experiments. O.M. and W.W.L. made the thermal stimulation device. W.W.L. performed the experiments. W.W.L. and R.I.W. analysed the data and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to R.I.W. (rachel_wilson@hms.harvard.edu).

METHODS

Fly stocks. The genotypes used in each figure are as follows:

Figure 1a–d: *GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*

Figure 1e–h: *GMR49B06-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP* and *GMR67D03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*.

Figure 2a–c: *GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*

Figure 2d: *Gr28b^{MB03888};GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*

Figure 2e, f: *LexAOp-HA-Ort/Gr28b.d-LexA;GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP/+*

Figure 2g: *GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP* (wild type) and *Gr28b^{MB03888};GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP* (*Gr28b.d^{-/-}*)

Figure 3a–c: *GMR54A03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*

Figure 3d: *Gr28b^{MB03888};GMR54A03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*

Figure 3e–g: *LexAOp-HA-Ort/Gr28b.d-LexA;GMR54A03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP/+*

Figure 4a–f: *pJFRC7-/+;Gad1-Gal4/+*

Figure 4g: *20XUAS-GCaMP3/+;Gr28b.d-Gal4/+* (warm cells) and *20XUAS-GCaMP3/+;GMR79C04-Gal4/+* (cool cells)

Figure 4h: *GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*

The Gal4 lines used to drive transgene expression in specific cell types are as follows: *GMR54A03-Gal4*: 1 warm-cool-PN

GMR49B06-Gal4: 1 slow-cool-PN

GMR67D03-Gal4: 1 slow-cool-PN (different from that labelled by *49B06*)

GMR95C02-Gal4: ~4 fast-cool-PN and 1 warm-PN

GMR52G03-Gal4: ~3 glutamatergic LNs (see Extended Data Fig. 10)

GMR91H10-Gal4: ~3 glutamatergic LNs (see Extended Data Fig. 10)

GMR79C04-Gal4: 3 cool thermoreceptor cells in the arista

Gr28b.d-Gal4: 3 warm thermoreceptor cells in the arista

Gad1-Gal4: multiple GABAergic LNs

Flies (*Drosophila melanogaster*) were raised on standard cornmeal agar medium supplemented with rehydrated potato flakes on a 12 h light/dark cycle at 25 °C. All experiments were performed on adult female flies between 6 h and 2 days post-eclosion. The *GMR Gal4* lines are described previously¹⁴ and were obtained from the Bloomington Drosophila Stock Center (BDRC). The *GMR* line that labels the warm-cool-PN (*54A03*) also stochastically labelled 1–2 additional nearby somata; we were not able to obtain recordings from these additional neurons due to weak and unreliable labelling. The *GMR* line that labels three cool thermoreceptor cells in the arista (*79C04*) was identified based on a visual screen for expression in the proximal antennal protocerebrum and was validated using *GCaMP3* imaging of the arista (Fig. 4g and Extended Data Fig. 6); it reliably drives expression of fluorescent reporter transgenes in only three cells in the arista, in contrast to a previously-identified line, *NP4486-Gal4* (ref. 2), which can drive expression in >3 cells in the arista on a stochastic basis (data not shown). *Gr28b.d-Gal4* was first published in ref. 13. *Gad1-Gal4* (line 2B) was constructed by Toshi Kitamoto using a *Gad1* promoter fragment and was first published in ref. 26, where it was called ‘*Gad2B-Gal4*’. *Gr28b^{MB03888}* was reported in ref. 27, and also in ref. 3 (where it was termed the ‘Mi’ allele); we obtained it from the BDRC. *pJFRC2-10XUAS-IVS-mCD8::GFP* (*su(Hw)attP1*) and *26XLexAop2-mCD8::GFP(attP2)* were reported in ref. 28 and were obtained from the BDRC. *20XUAS-GCaMP3 (attP18)* was reported in ref. 29 and obtained from the BDRC. *pJFRC7-20XUAS-IVS-mCD8::GFP (attP40)* was reported in ref. 30; before that report, the same transgene was described in another insertion site²⁸. *UAS-nls:GFP* was obtained from the BDRC. *LexAOp-HA-Ort* was a generous gift from C. -H. Lee.

Generation of transgenic flies. The *Gr28b.d* promoter fragment was PCR amplified from genomic DNA and assembled into the *EcoRI/KpnI* digested pBPnlsLexA::GADflUw vector (Addgene) using the Gibson Assembly Master Mix (NEB). The PCR primers used were 5′-GGCCCTTTCGTCTTCAAGAATTCGTCACCATGCCCTGCACTTCACC-3′ and 5′-GGGTGGCATTTTGATTGCTAGCGGTA CCTCAACGTATTTTAAATATTTATAGTATGTA-3′. The underlined regions indicate regions of overlap with the vector. Constructs were sequenced before injection into *y,w;attP40* embryos. Genetic Services, Inc. provided transgenic production services. The expression patterns of transgenic flies were verified by crossing each *LexA* line to a GFP reporter and examining GFP expression in the arista and the proximal antennal protocerebrum (Extended Data Fig. 6).

The *LexAOp-HA-Ort* vector was constructed by inserting the *2xHA-ORT* fragment from *pUAST-2xHA-ORT* (ref 31) into the *pLOT* vector³², and it was injected in *y,w* embryos via P-element transformation.

Electrophysiological recordings. *In vivo* whole-cell patch clamp recordings from neural somata were performed under visual control using an upright compound microscope to direct the electrode to a GFP-positive soma. The fly was restrained in a platform which allowed the brain and antennae to be bathed in saline while most of the fly remained dry. Recordings were performed essentially as previously described²⁵, except for the following modifications detailed in ref 20. The head was

rotated 180° around the neck connective, so that the ventral side of the brain was facing upwards and therefore accessible to visualization via the water-immersion objective above the preparation. The fly remained alive throughout the experiment even when the head was rotated in this manner, as evidenced by continual spontaneous movements of the body, as well as normal brain activity. It was necessary to rotate the head because both the PNs and the LNs that innervate the proximal antennal protocerebrum have somata ventral to the antennal lobes, and so are inaccessible when the brain is dorsal-side up.

The brain was perfused in external saline containing 103 mM NaCl, 3 mM KCl, 5 mM *N*-tris(hydroxymethyl) methyl-2-aminoethane-sulfonic acid, 8 mM trehalose, 10 mM glucose, 26 mM NaHCO₃, 1 mM NaH₂PO₄, 1.5 mM CaCl₂, and 4 mM MgCl₂ (osmolality adjusted to 270–275 mOsm). The saline was bubbled with 95% O₂/5% CO₂ to a pH of 7.3. The internal solution for patch-clamp pipettes were contained the following: 140 mM potassium aspartate, 10 mM HEPES, 1 EGTA, 4 mM MgATP, 0.5 mM Na₃GTP, 1 mM KCl, and 13 mM biocytin hydrazide. The pH of the internal solution was adjusted to 7.2 and the osmolality was adjusted to ~265 mOsm.

In initial experiments, we filled all recorded neurons with biocytin to visualize their morphology. Some of the lines we used in this study drive Gal4 expression in more than one neuron in the region where the somata of thermosensory PNs are located (just ventral to the antennal lobe), but we found that each morphological type of neuron had characteristic electrophysiological properties and temperature responses, and so we could accurately identify each type within the relevant Gal4 line based on these characteristics alone. Hence in subsequent recordings, neurons were classified based on these criteria and not every cell was filled. However, PN recordings made in the *Gr28b.d* mutant background were always filled, and LNs were also always filled.

Recordings were performed with an Axopatch 200B amplifier (Axon Instruments). Recorded voltages were low-pass filtered at 5 kHz and digitized at 10 kHz. The seal conductance can create a small tonic depolarization in small neurons³³ so in many experiments we injected a small amount of constant hyperpolarizing current to bring the cell back down to its native resting potential. We estimated the native resting potential of all these cell types to be –40 mV, based on measuring spontaneous spiking in cell-attached mode before rupturing the seal, and then matching the spontaneous spike rate in whole-cell mode³³. Spontaneous spikes are typically not visible in cell-attached recordings from warm-PNs, and so we assumed a resting potential of –40 mV for these cells. In most experiments where a cell was depolarized as a result of blocking synaptic inhibition, we returned the cell to its original potential before measuring its responses to thermal stimuli, so as to avoid any confounding effects of depolarization per se. All reported membrane potential values are uncorrected for the liquid junction potential, which is –13 mV (ref 33).

In some experiments, we used picrotoxin (100 μM) and CGP54626 (50 μM) to block synaptic inhibition mid-way through the experiment. These drugs were added to the saline perfusate. In most experiments where a cell was depolarized as a result of blocking synaptic inhibition, we returned the cell to its original potential before measuring its responses to thermal stimuli. This procedure minimized any confounding effects of depolarization per se. In warm-cool-PNs, blocking synaptic inhibition tended to induce oscillations in the membrane potential, and so small effects of this manipulation should be interpreted with caution.

LN recordings. Flies with *Gad1-Gal4* driving *pJFRC7-20XUAS-IVS-mCD8::GFP* were analysed with dual confocal immunofluorescence microscopy, using anti-CD8 and anti-GABA antibodies, as previously described²⁵. We found that 90% of CD8::GFP-positive cells in the cluster of cell bodies just ventral to the antennal lobe were also GABA-positive. Because we restricted our electrodes to cell bodies in that region, and we also filled each recorded cell with biocytin to verify that it innervated the proximal antennal protocerebrum, it is likely that almost all the LNs we recorded using this Gal4 line are GABAergic. In these experiments, around two-thirds of cells that we recorded from in the *Gad1-Gal4* line did not respond to our thermal stimuli and were discarded; many of these are likely to represent antennal lobe PNs and LNs, based on the morphologies we observed when we filled a subset of these cells with biocytin. We used a similar procedure to target our electrodes to the glutamatergic LNs: flies with either *GMR52G03-Gal4* or *GMR91H10-Gal4* driving *pJFRC2-10XUAS-IVS-mCD8::GFP* were immunostained using anti-CD8 and anti-dVGluT antibodies, as previously described²⁰. We found that all the CD8::GFP-positive cells in the cluster of cell bodies just ventral to the antennal lobe were dVGluT-positive. As before, we restricted our electrodes to cell bodies in that region, and we filled each recorded cell with biocytin to verify that it arbourized in the proximal antennal protocerebrum.

Thermal stimulus delivery. Because all neurons are sensitive to temperature, and because temperature gradients between the recording and ground electrodes can produce electrical artefacts, we wanted to spatially restrict temperature changes to the peripheral thermoreceptor cells in the antenna to specifically study how peripheral signals are processed in the brain. Therefore, we designed a jet micro-thermode

capable of delivering a fast temperature stimulus to a $\sim 150\ \mu\text{m}$ target (Extended Data Fig. 1). The thermode system consists of two components: a room-temperature, laminar flow saline perfusion chamber (Warner Instruments, RC-26GLP), and a probe that delivers a jet of temperature-controlled saline. The fly is placed at the centre of the chamber and the probe is positioned $\sim 100\ \mu\text{m}$ from the ipsilateral antenna, oriented so that the jet of saline flows across both arista and both antennae but not the rest of the fly. The tip of the arista is immobilized with light-curing adhesive. The laminar flow of the perfusion chamber ensures that the saline jet continues on past the antenna and out of the chamber without substantial mixing with the rest of the bath. The bath temperature was kept at $23\text{--}25\ ^\circ\text{C}$. We verified that the baseline firing rate of all PN types adapted almost completely within a few minutes to temperature changes within this range. The hydrodynamics of the system were visualized by adding a solution of $500\ \mu\text{M}$ phenol red to the saline jet. We adjusted the flow rates of the perfusion chamber and saline jet until no visible mixing occurred. Simultaneous thermocouple measurements (see Temperature Measurements, below) were performed to ensure that the temperature of the bath near the brain remained stable when temperature stimuli were presented to the antenna. The intrinsically thermosensitive neurons of the brain¹² would therefore be insensitive to our stimuli. As additional controls for the spatial specificity of our thermal stimuli, we verified that cutting the antennal nerves bilaterally abolished responses in all PN types. Additionally, removing the arista bilaterally essentially eliminated all temperature responses in the fast-cool-PNs (see Fig. 1b), and it completely eliminated the responses of the warm-PNs, warm-cool-PNs, and glutamatergic LNs (data not shown).

The jet was produced by sending saline from a reservoir (pressurized to $\sim 5\ \text{p.s.i.}$) through a narrow tube that terminated in a $160\ \mu\text{m}$ nozzle (Small Parts, 0.0063" ID polyimide tubing, TWPT-0063-30). Upstream of the nozzle, the tube passed through a heat exchanger made from two concentric pipes. The outer pipe was fed with a constant flow of water pumped in from a hot bath ($\sim 40\ ^\circ\text{C}$) or a cold bath ($\sim 0\ ^\circ\text{C}$). Water from only one of the baths flowed into the heat exchanger at a time. A system of three three-way solenoid valves selected between hot mode and cold mode.

The heat exchanger was powerful enough to bring the saline jet (as measured at the antenna) to $\sim 13\ ^\circ\text{C}$ in cold mode and $\sim 34\ ^\circ\text{C}$ in hot mode. After switching modes, the jet temperature took $\sim 7\ \text{s}$ to plateau. During this slow transition, the valves could also stop flow entirely to the outer pipe (no-flow mode) when the heat exchanger was at room temperature ($\sim 24\ ^\circ\text{C}$). By switching between these three modes, we could achieve slow temperature transitions between three different temperatures: $\sim 13\ ^\circ\text{C}$, $\sim 24\ ^\circ\text{C}$ and $\sim 34\ ^\circ\text{C}$. This method was the basis for our slow cool and slow warm stimuli.

Finally, a separate 'fast' solenoid valve was installed in line with the pressurized saline tube. Opening and closing this fast valve turned the jet on or off. When the jet is turned off, the steady laminar flow of the bath quickly washes over the antenna, bringing it back to bath temperature ($\sim 24\ ^\circ\text{C}$). Switching this valve allowed for fast ($\sim 100\ \text{ms}$) transitions between a hot (or cold) jet and the $\sim 24\ ^\circ\text{C}$ bath. This method was the basis for our fast cool and fast warm stimuli.

To produce small temperature steps, the heat exchanger was heated to $\sim 34\ ^\circ\text{C}$, then cooled with ice so that the jet delivered saline that was only a few degrees warmer than the bath. We then gradually cooled the heat exchanger with ice while periodically opening and closing the fast valve to produce a fast temperature change at the antenna (within $100\ \text{ms}$) to a temperature just above $24\ ^\circ\text{C}$. Over several minutes, the temperature in the heat exchanger progressively fell to $24\ ^\circ\text{C}$, and then continued to fall. In this manner, we were able to finely sample a range of temperatures several degrees above and below $24\ ^\circ\text{C}$. Because the stimulus control was performed manually in this case, we could not achieve precisely the same temperature stimuli in every experiment. We therefore recorded the temperature at the antenna in each stimulus trial and post hoc binned the data by temperature to construct the sensitivity plots showing average responses to small temperature steps. In the course of these experiments, we observed that central neurons did not respond when the jet temperature precisely equalled the bath temperature (to the limits of precision of our thermocouple), indicating that these neurons do not respond to any mechanical effects of the jet.

Temperature measurements. Temperature measurements at the location of the antenna were made using K-type thermocouples with a $75\text{--}100\ \mu\text{m}$ tip (Omega Engineering, CHAL-001 thermocouple and SMPW-CC-K-M connector). One wire of the thermocouple was threaded into a microbore Teflon tube (Cole-Parmer, EW-06417-11) to provide mechanical rigidity and insulation between the two thermocouple wires. The thermocouple was mounted on a micromanipulator and its tip was positioned entirely within the saline jet, which included $500\ \mu\text{M}$ phenol red for visualization. Thermocouple measurements were acquired and digitized at $99\ \text{Hz}$ by a National Instruments USB-9213 Thermocouple Measurement Device. Temperature measurements displayed in the figures represent the mean of several experiments with the same commands to the stimulus device; because these measurements

were highly consistent across experiments (generally within $1\ ^\circ\text{C}$ of the mean) we did not measure the temperature in every experimental replicate.

Histochemistry. In some experiments (noted above), the morphology of the recorded neuron was visualized after recording by incubating the brain with a fluorescent conjugate of streptavidin, as published previously³⁴. Immunohistochemistry was performed as described previously²⁵. Primary antibodies were obtained from the following sources (with dilutions in parentheses): mouse nc82 from the Developmental Studies Hybridoma Bank³⁵ (nc82-s, 1:50), rat anti-CD8 from Invitrogen (MCD0800, 1:200), rabbit anti-GABA from Sigma²⁵ (A2052, 1:200), rabbit anti-dVGlut (1:500; gift of A. DiAntonio, ref. 19). Secondary antibodies (Invitrogen) were used at a 1:250 dilution. To reconstruct neuronal morphology from biocytin fills, we hand-traced the skeletonized morphology using the Simple Neurite Tracer plugin in Fiji, using the Fill Out command to automatically generate a 3D volume, which we subsequently converted to a z projection. The morphology of the slow-cool-PN has been described previously³⁶.

Calcium imaging. Cool and warm cells were imaged in separate experiments. Widefield illumination was generated with a $100\ \text{W}$ mercury arc lamp (Olympus) attenuated with a ND-25 neutral-density filter, band-pass filtered at $540\text{--}580\ \text{nm}$, and delivered to the specimen through a $40\times$ water-immersion objective mounted in an upright compound microscope (Olympus BX51WIF). Phenol red was omitted from the saline jet during these experiments. Images were acquired using a CMOS camera (Hamamatsu ORCA-Flash4.0 V2) with a $50\ \text{ms}$ exposure time. Thermal stimuli were delivered in fast mode (that is, with a rise time of $100\ \text{ms}$ or less). Data was analysed using custom MATLAB software. Calcium imaging confirmed that *Gr28b.d-Gal4* drives expression in warm thermoreceptor cells in the arista, whereas *GMR79C04-Gal4* drives expression in cool thermoreceptor cells in the arista (Extended Data Fig. 6).

Histamine iontophoresis. For histamine iontophoresis, a high-resistance ($\sim 100\ \text{M}\Omega$) sharp pulled glass micropipette was filled with a solution of $1\ \text{M}$ histamine dihydrochloride in saline. The pipette was inserted into the ipsilateral third antennal segment, and histamine was ejected using a $0.5\text{--}3\ \text{s}$ pulse of positive current ($300\ \text{nA}$) applied with an iontophoresis current generator (Model 260, World Precision Instruments) gated by a TTL pulse. Histamine ejection was initiated $4\text{--}5\ \text{s}$ before the onset of the thermal stimulus. A constant negative backing current of $40\ \text{nA}$ was applied to retain histamine in the pipette between ejections. These ejection parameters were chosen such that there was essentially no effect of histamine iontophoresis in control genotypes that lacked the LexA transgene (Extended Data Fig. 7), and the brief electrical artefact produced by the iontophoresis current pulse was small. Although some thermosensory PNs arborize bilaterally, in pilot experiments where we removed the contralateral antenna we observed no clear alteration in the thermal responses of these cells, and so we performed histamine iontophoresis only in the ipsilateral antenna. We chose to iontophorese histamine into the antenna in these experiments because in pilot studies we found that iontophorese histamine into the brain had effects on the central neurons that did not depend on the *LexAOp-HA-Ort* transgene.

We also created a LexA line designed to specifically drive expression in the cool thermoreceptor cells by converting the *GMR79C04-Gal4* line to a LexA line using established methods²⁸, and we used this to drive expression of the histamine receptor in the cool thermoreceptor cells. However, when we attempted to silence the cool cells in this manner, we did not substantially alter the thermal responses of fast-cool-PNs, which clearly receive almost all their thermal input from the arista. The failure of this positive control is likely due to low transgene expression levels in the cool thermoreceptor cells, because GFP signals driven by this LexA line were also weak.

Data analysis. Reported numbers of replicates (n values) always represent numbers of cells recorded; in almost all cases, only one cell was recorded per brain. Centre values represent the mean and error bars represent s.e.m. computed across cells (experiments). No formal statistical calculations were used to pre-determine sample sizes. Sample sizes were qualitatively governed by the observed cell-to-cell reliability in the measurements we performed in pilot experiments. Our sample sizes are similar to those generally employed in the field. Data distribution was assumed to be normal with similar variance between the groups that were compared statistically, but this was not formally tested. All statistical tests were two-sided. In no case were flies from the same genotype assigned to distinct experimental treatments, and so no randomization procedure was needed. No blinding was performed during experiments or analysis.

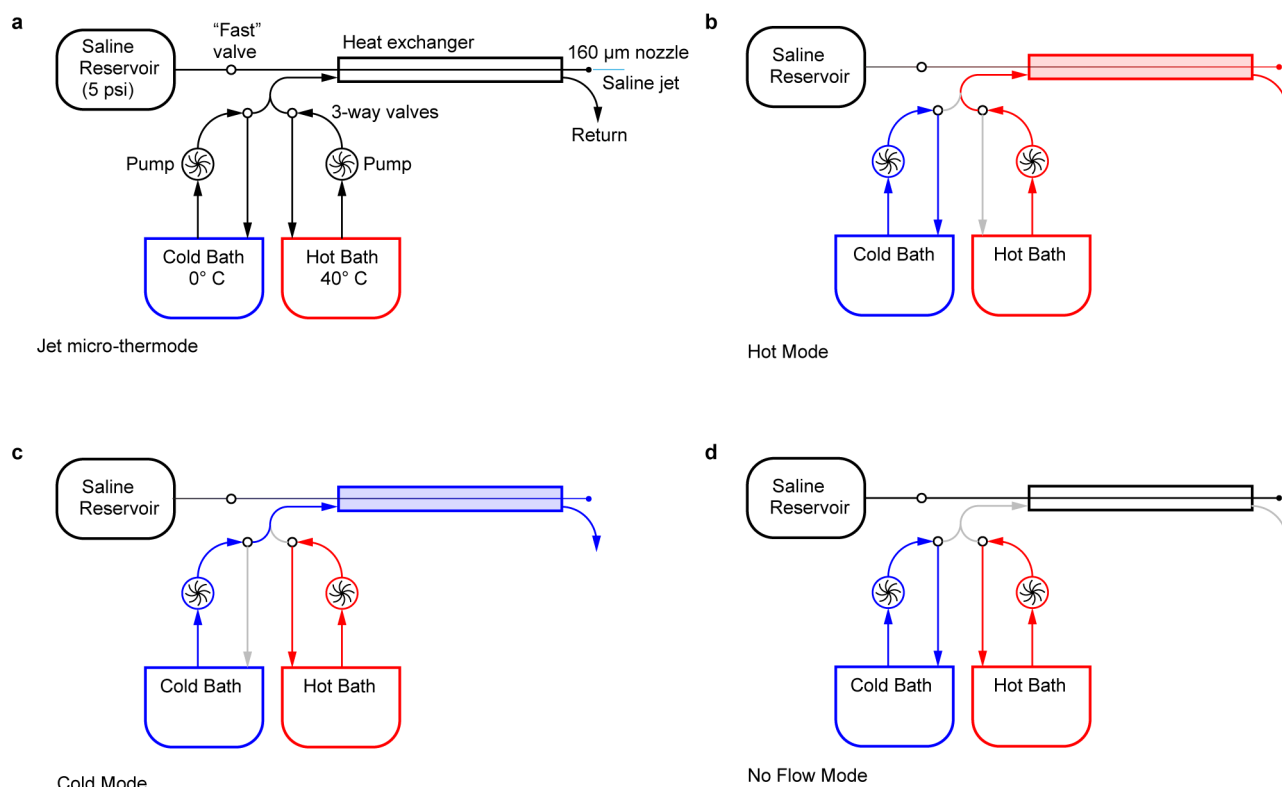
Spikes were detected using custom software in Igor Pro. Peri-stimulus time histograms were generated by calculating the firing rate in $50\ \text{ms}$ bins that overlapped by $25\ \text{ms}$. Data for each cell represents an average of at least 4 trials of the same stimulus for electrophysiology experiments, and an average of 3 trials for calcium imaging experiments. All error bars represent s.e.m. computed across cells, as do the shaded bands in the peri-stimulus time histograms. For calcium imaging experiments, fluorescence was measured within a region of interest near the base of the

arista that was drawn so as to contain the somata of 1–2 GCaMP-expressing cells in the same focal plane.

To compare responses to large temperature steps, we counted spikes in a 2 s period after step onset. To compare the sensitivity of cells to small temperature steps, we measured both mean temperature and mean neural activity (either firing rate or membrane potential) in a time window 100–400 ms after opening the valve for all cells except the warm-cool-PN, where we used a window of 100–200 ms after valve opening. Where we measured the mean membrane potential within a time window, we first low-pass filtered the membrane potential at 10 Hz to remove spikes. Because our device does not permit forward control of temperature with sub-degree precision, the exact temperatures we used were slightly different in each experiment. We therefore binned the data for the sensitivity analyses according to the measured temperature in each trial, averaging together different measurements in the same temperature bin within the same experiment. For each temperature bin we then computed then mean neural activity across experiments, and also the s.e.m. across experiments to generate the sensitivity plots; the *x*-axis value of each data point in the sensitivity plots represents the middle of each temperature bin. In order to perform tests of statistical significance on the sensitivity data, we used an iterative approach to avoid a multiple comparisons error. There is good reason to think there might be a true difference for small temperature steps that would not be evident for larger temperature steps, but we did not know a priori what temperature changes these neurons might be sensitive to. Therefore, for a given pair of genotypes and a given stimulus polarity (warming or cooling), we first compared responses to the smallest temperature step (0.125 °C) using an unpaired *t*-test. We then adjusted the threshold for significance (α) to reflect the cumulative number of tests performed (*n*) using a Bonferroni correction ($\alpha = 0.05/n$) and proceeded to test the adjacent

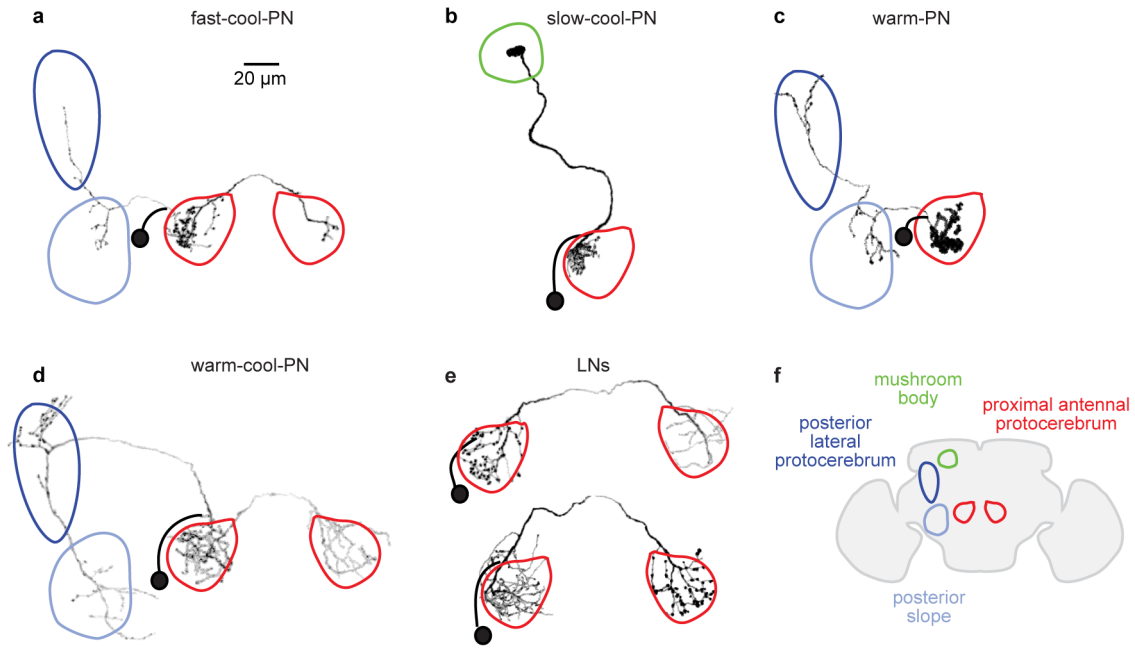
temperature step (0.425 °C), and so on. This procedure maximized our statistical power for the smallest steps. Because statistical power decreased for increasingly larger steps, this procedure is appropriate for identifying the smallest step where a difference might occur, but not for identifying the entire temperature range over which there might be a difference.

26. Mehren, J. E. & Griffith, L. C. Cholinergic neurons mediate CaMKII-dependent enhancement of courtship suppression. *Learn. Mem.* **13**, 686–689 (2006).
27. Xiang, Y. *et al.* Light-avoidance-mediating photoreceptors tile the *Drosophila* larval body wall. *Nature* **468**, 921–926 (2010).
28. Pfeiffer, B. D. *et al.* Refinement of tools for targeted gene expression in *Drosophila*. *Genetics* **186**, 735–755 (2010).
29. Tian, L. *et al.* Imaging neural activity in worms, flies and mice with improved GCaMP calcium indicators. *Nature Methods* **6**, 875–881 (2009).
30. Fisek, M. & Wilson, R. I. Stereotyped connectivity and computations in higher-order olfactory neurons. *Nature Neurosci.* **17**, 280–288 (2014).
31. Takemura, S. Y. *et al.* Cholinergic circuits integrate neighboring visual signals in a *Drosophila* motion detection pathway. *Curr. Biol.* **21**, 2077–2084 (2011).
32. Lai, S. L. & Lee, T. Genetic mosaic with dual binary transcriptional systems in *Drosophila*. *Nature Neurosci.* **9**, 703–709 (2006).
33. Gouwens, N. W. & Wilson, R. I. Signal propagation in *Drosophila* central neurons. *J. Neurosci.* **29**, 6239–6249 (2009).
34. Wilson, R. I., Turner, G. C. & Laurent, G. Transformation of olfactory representations in the *Drosophila* antennal lobe. *Science* **303**, 366–370 (2004).
35. Laissue, P. P. *et al.* Three-dimensional reconstruction of the antennal lobe in *Drosophila melanogaster*. *J. Comp. Neurol.* **405**, 543–552 (1999).
36. Jenett, A. *et al.* A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2**, 991–1001 (2012).



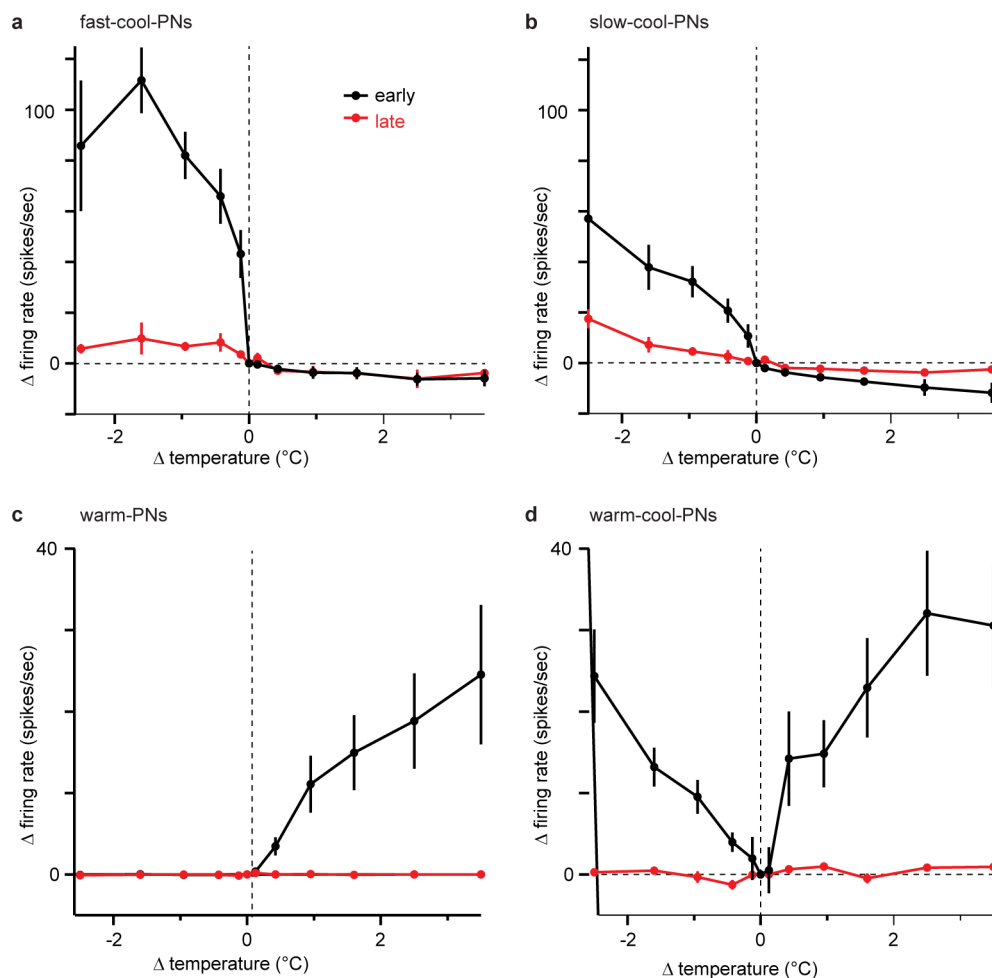
Extended Data Figure 1 | Jet micro-thermode schematic. **a**, Schematic showing the major components of the jet micro-thermode. A pressurized reservoir forces saline through the inner chamber of a heat exchanger. The saline is then expelled as a jet through a 160 μm nozzle. This jet can be quickly turned on and off via a two-way solenoid valve (the 'fast' valve). The outer chamber of the heat exchanger is fed by two pumps which circulate water from a hot and a cold water bath. Two three-way solenoid valves control the flow of

water into the exchanger. A third valve (not shown) serves to return the outflow of the heat exchanger to the same bath that is feeding it. **b**, **c**, In hot mode and cold mode the heat exchanger is heated or cooled by the continual circulation of water from one of the two baths through the outer chamber. **d**, In the no-flow configuration, circulation in the outer chamber is stopped. This keeps the heat exchanger at a fairly constant temperature, stabilized by the thermal mass of the water in the outer chamber.



Extended Data Figure 2 | Morphologies of neurons in the proximal antennal protocerebrum. Images are *z* projections of coronal sections (dorsal is up). The soma remains sealed to the patch pipette at the end of each recording, and so the soma and the proximal part of the primary neurite are not present in the reconstructed morphologies; they are indicated schematically (soma with a black ball and primary neurite with a thick black line). Morphologies shown here were representative of all filled examples of the same neuron type (except as noted below). Regions are colour-coded according to the key in **f**. **a**, A fast-cool-PN. These neurons have dendrites in the proximal antennal protocerebrum and axons projecting to two regions of the lateral protocerebrum (posterior lateral protocerebrum and posterior slope). Some fast-cool-PNs arbourized unilaterally (rather than bilaterally) in the proximal antennal protocerebrum; we pooled data from both morphologies because their physiological properties were indistinguishable. **b**, A slow-cool-PN. These neurons have dendrites in the ipsilateral proximal antennal protocerebrum and axons projecting to a subregion of the calyx of the mushroom body. This

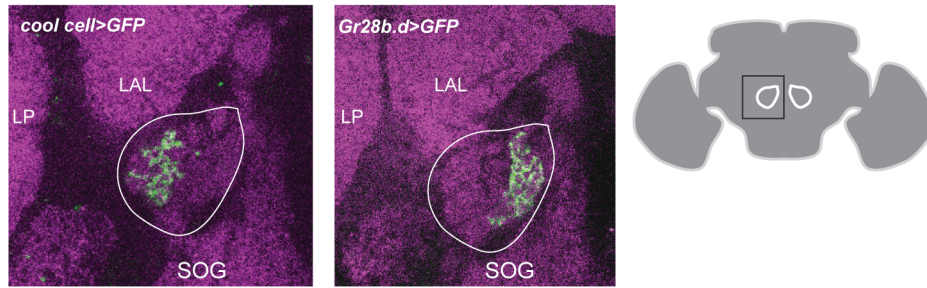
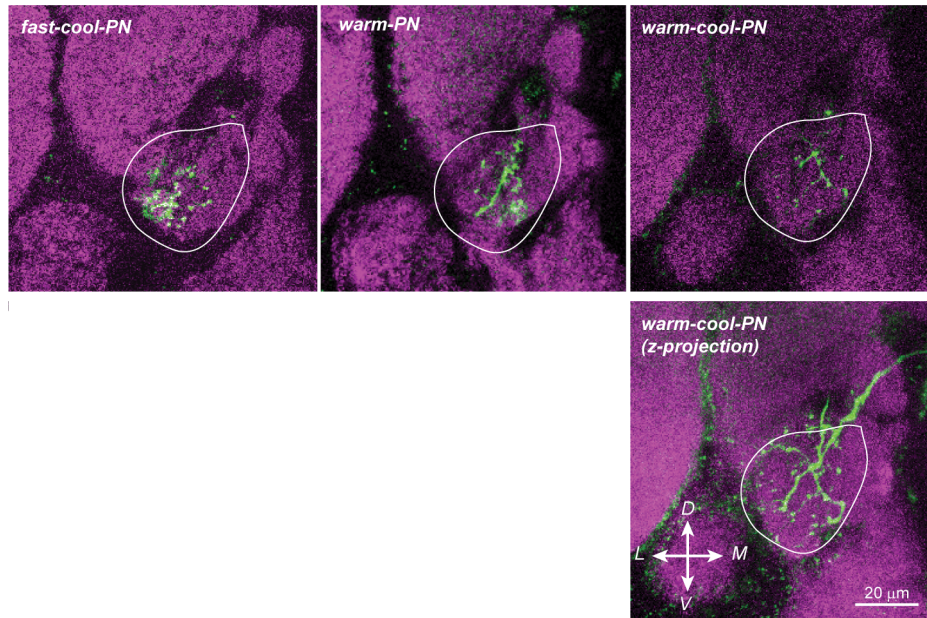
morphology has been described previously³⁶ (that study called this neuron a 'transverse antennal lobe projection neuron'). **c**, A warm-PN. These neurons have dendrites in the ipsilateral proximal antennal protocerebrum and axons projecting to two regions of the lateral protocerebrum (posterior lateral protocerebrum and posterior slope). **d**, A warm-cool-PN. These neurons have dendrites in both the ipsi- and contralateral proximal antennal protocerebrum, as well as the ventral-posterior margin of the antennal lobe. Their axons project to two regions of the lateral protocerebrum (posterior lateral protocerebrum and posterior slope). **e**, Two GABAergic LNs. All LNs we encountered in this study were bilateral. We define an LN here as a neuron that does not extend processes into any region except the proximal antennal protocerebrum (and the midline commissure that links the right and left copies of this brain region). Most LNs arbourized broadly within the posterior antennal protocerebrum and so are well-positioned to mediate cross-talk between warm and cool pathways. **f**, Locations of colour-coded brain regions in a coronal section through the brain.



Extended Data Figure 3 | Adaptation to small temperature steps.

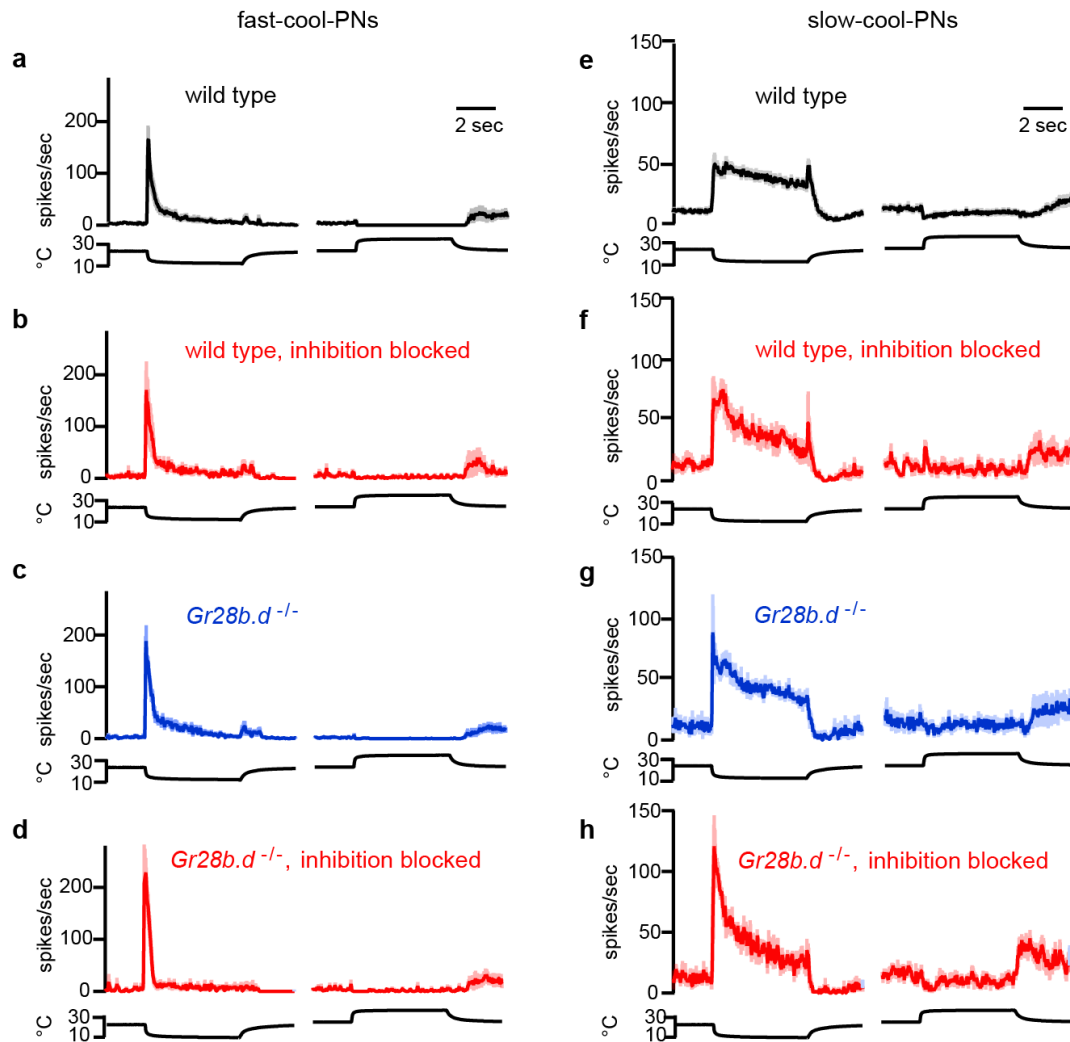
a–d, Sensitivity to small temperature steps, shown as the mean change in firing rate versus the change in temperature, averaged across experiments, \pm s.e.m. ($n = 5$ – 9). Data in red are measured during a late period after stimulus onset (from 4 to 5 s after valve opening). Data in black are measured during an early period after stimulus onset (from 100 to 400 ms after valve opening,

except for the warm-cool-PNs, where it was from 100 to 200 ms). Data in black are reproduced from Fig. 1c (**a**, fast-cool-PNs), Fig. 1g (**b**, slow-cool-PNs), Fig. 2g (**c**, warm-PNs) and Extended Data Fig. 8 (**d**, warm-cool-PNs). Note that slow-cool-PNs show little adaptation to large temperature steps (Fig. 1g). However, their responses to small steps adapt strongly, like those of the other PN types.

a axons of peripheral neurons**b** dendrites of central neurons

Extended Data Figure 4 | Innervation patterns in the proximal antennal protocerebrum. Images are single coronal confocal sections through the proximal antennal protocerebrum, except where noted below. Magenta is neuropil (nc82 immunofluorescence). The approximate boundary of the proximal antennal protocerebrum is indicated in white outline. **a**, Axon terminals of cold peripheral neurons and warm peripheral neurons. GFP expression is driven by specific Gal4 lines (*GMR79C04-Gal4* for the cool cells, *Gr28b.d-Gal4* for the warm cells). As noted previously², the cool peripheral cells project to the lateral part of this brain region, whereas the warm peripheral cells project more medially. Neuropils surrounding the proximal antennal protocerebrum are labelled as LP (lateral protocerebrum), LAL (lateral accessory lobe) and SOG (subesophageal ganglion). The schematic on the right

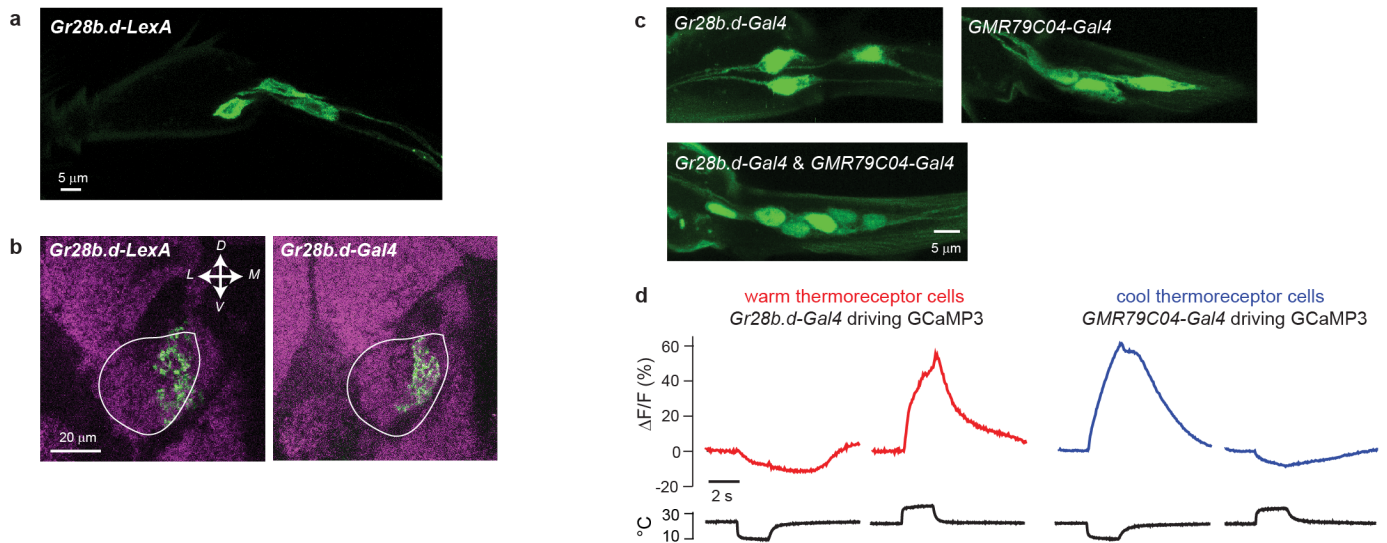
shows a coronal section of the whole brain with the location of the proximal antennal protocerebrum indicated by white outlines. **b**, Dendrites of three types of projection neurons. Each cell was filled with biocytin and visualized using a fluorescent streptavidin conjugate. The dendrites of the fast-cool-PNs reside in the lateral part of the proximal antennal protocerebrum. The same is true of the dendrites of the slow-cool-PNs (data not shown). The dendrites of the warm-PNs overlap with the axons of both warm and cool peripheral neurons. The dendrites of warm-cool-PNs are sparser than those of other PN types, so in a single section only scattered fragments of dendrite are visible; therefore a z projection through the entire proximal antennal protocerebrum is shown in the image below.



Extended Data Figure 5 | Circuit contributions to cool-PN responses.

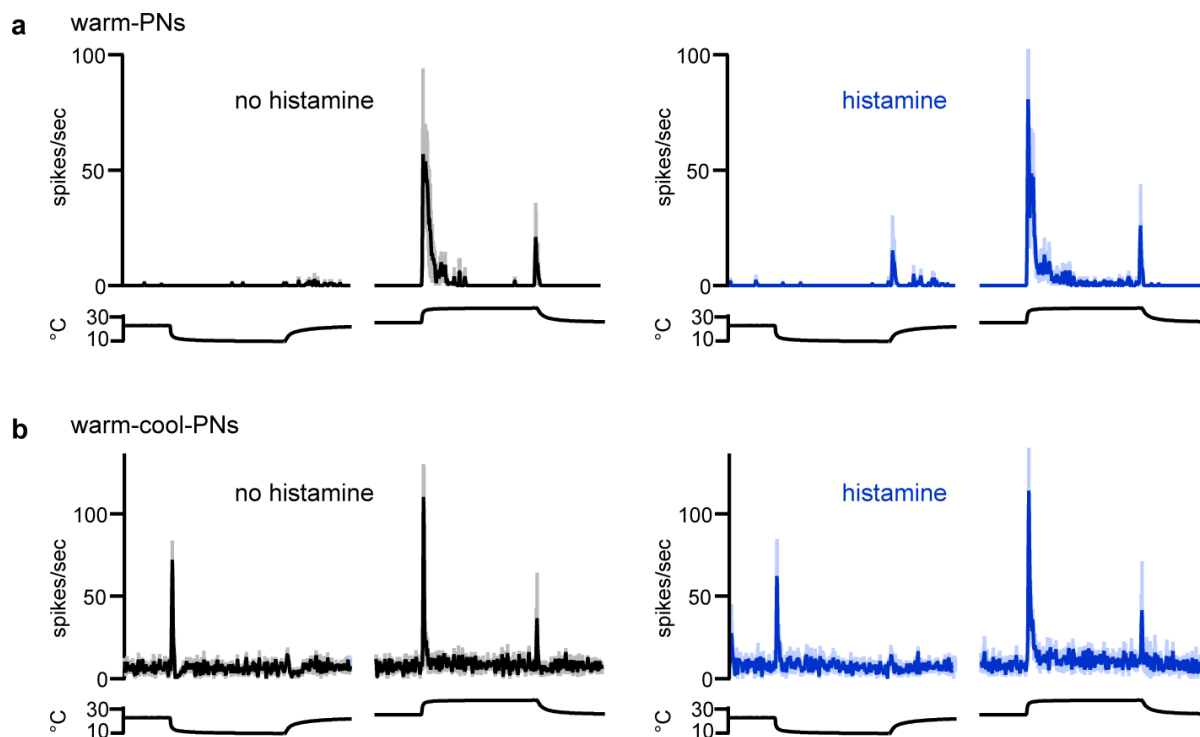
a–d, Fast-cool-PNs. **a**, Mean firing rate for the fast-cool-PNs, \pm s.e.m. Reproduced from Fig. 1b. **b**, Mean firing rate with synaptic inhibition blocked with picrotoxin and CGP54626 ($n = 4–5$). **c**, Mean firing rate in the *Gr28b.d* mutant ($n = 5–10$). **d**, Mean firing rate in the *Gr28b.d* mutant with inhibition blocked ($n = 4–6$). **e–h**, Slow-cool-PNs. **e**, Mean firing rate for the slow-cool-PNs \pm s.e.m. Reproduced from Fig. 1f. **f**, Mean firing rate with synaptic inhibition blocked with picrotoxin and CGP54626 ($n = 4–5$). **g**, Mean firing

rate in the *Gr28b.d* mutant ($n = 4–5$). **h**, Mean firing rate in the *Gr28b.d* mutant with inhibition blocked ($n = 4$). In the *Gr28b.d* mutant, note that blocking inhibition produces only modest disinhibition in the cool-PNs. This contrasts with our results in the warm-PNs and warm-cool-PNs, where blocking inhibition in the mutant abolished all excitation evoked by the preferred stimulus for these neurons (Fig. 2 and Fig. 3). Genotypes are: **a–d**, *Gr28b*^{MB03888}; *GMR95C02-Gal4*; *pJFRC2-10XUAS-IVS-mCD8::GFP* and **e–h**, *Gr28b*^{MB03888}; *GMR67D03-Gal4*; *pJFRC2-10XUAS-IVS-mCD8::GFP*.



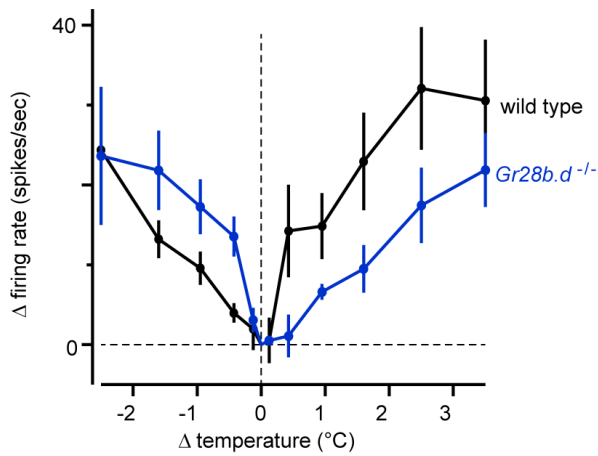
Extended Data Figure 6 | Validation of the LexA line for warm peripheral cells and the Gal4 line for cool peripheral cells. There are six thermoreceptor cells at the base of the arista, comprising three warm cells and three cool cells². *Gr28b.d-Gal4* labels the three warm cells³. **a**, **b**, Validation of the *Gr28b.d-LexA* line for warm peripheral cells. **a**, We generated a LexA line using a *Gr28b.d* promoter fragment. Here we show that *Gr28b.d-LexA* labels three cells at the base of the arista. This image is a z projection of a confocal stack through the base of the arista. Genotype is *Gr28b.d-LexA/+;26XLexAop2-mCD8::GFP/+*. **b**, Axon terminals in the brain of peripheral neurons labelled by *Gr28b.d-LexA* (left) and *Gr28b.d-Gal4* (right), both crossed with appropriate CD8::GFP reporter lines. As expected, the medial portion of the proximal antennal protocerebrum (outlined in white) is labelled with GFP. Images are single coronal confocal sections through the proximal antennal protocerebrum. Magenta is neuropil (nc82 immunofluorescence). Image on the right is reproduced from Extended Data Fig. 4. **c**, **d**, Validation of the *GMR79C04-Gal4*

line for cool peripheral cells. **c**, *GMR79C04-Gal4* labels three cells at the base of the arista. Combining this driver with *Gr28b.d-Gal4* labels six cells, indicating that these drivers label mutually exclusive populations of arista neurons. These images show z projections of confocal stacks through the base of the arista. Genotypes are *Gr28b.d-Gal4/+;UAS-nls-GFP/+* (left), *UAS-nls-GFP/+;GMR79C04-Gal4/+* (middle), *Gr28b.d-Gal4/+;UAS-nls-GFP/GMR79C04-Gal4* (right). **d**, Calcium imaging of neurons at the base of the arista shows that *Gr28b.d-Gal4* labels cells that are excited by warming and inhibited by cooling, whereas *GMR79C04-Gal4* labels cells that are excited by cooling and inhibited by warming. Shown here are representative experiments using the 'large fast step' stimuli. Genotypes are *20xUAS-GCaMP3/+;Gr28b.d-Gal4/+* (warm cell imaging) and *20xUAS-GCaMP3/+;GMR79C04-Gal4/+* (cool cell imaging). Note that Fig. 4g shows data for all experiments using 'small fast step' stimuli.

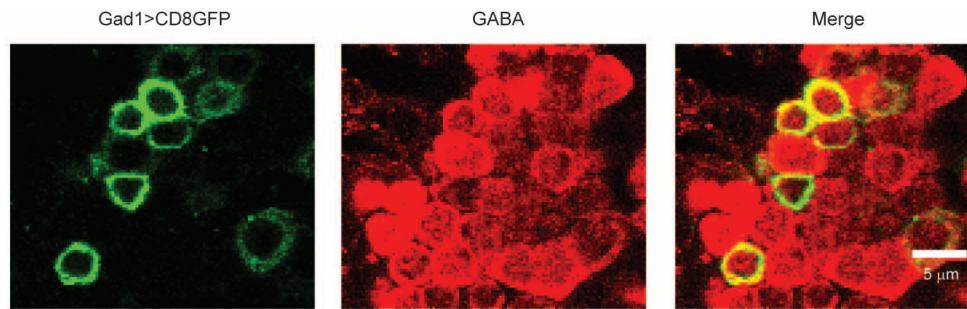


Extended Data Figure 7 | Histamine injection controls. In these experiments the LexA driver (*Gr28b.d-LexA*) was omitted, but otherwise the experimental protocol was the same as in Fig. 2e, f and Fig. 3e–g. After histamine was injected into the antenna, there was little effect on the thermal responses of

warm-PNs or warm-cool-PNs. **a**, Recordings from warm-PNs ($n = 5$). Genotype is *LexAOp-HA-Ort/+;GMR95C02-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*. **b**, Recordings from warm-cool-PNs ($n = 4$). Genotype is *LexAOp-HA-Ort/+;GMR54A03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*.

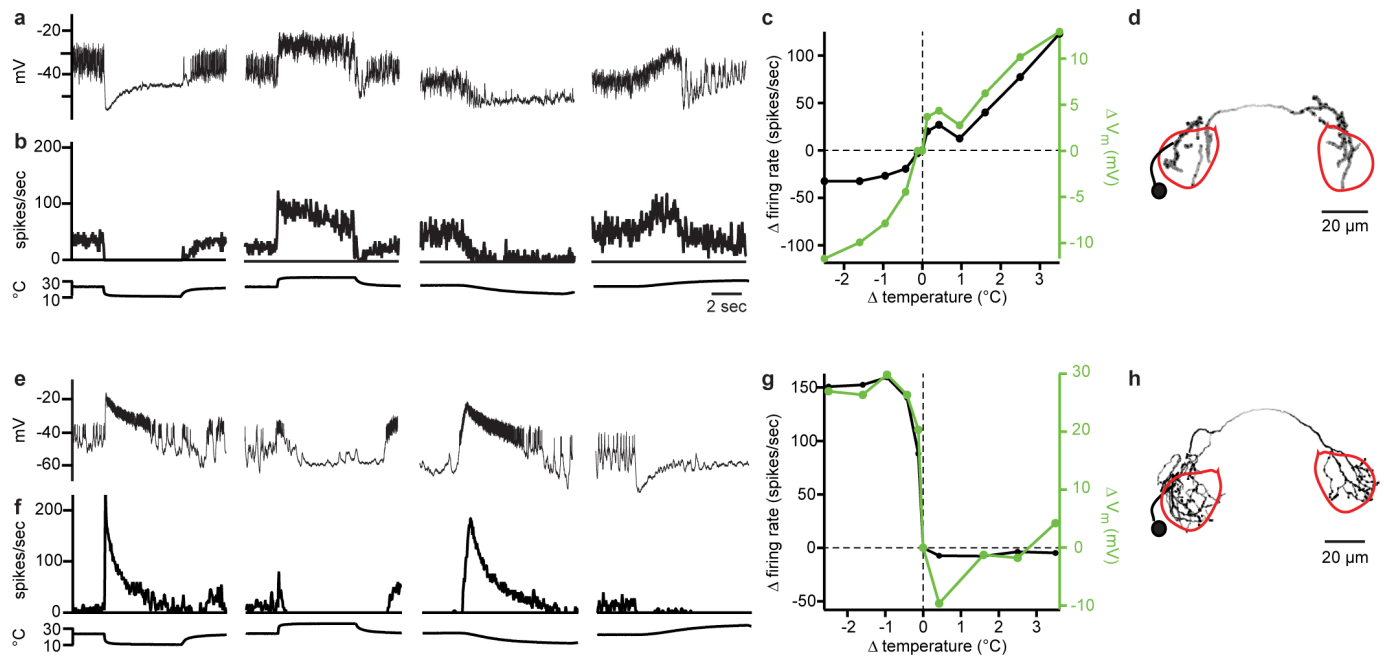


Extended Data Figure 8 | Sensitivity to small temperature steps in warm-cool-PNs. Sensitivity to small temperature steps, shown as the mean change in firing rate versus the change in temperature, averaged across experiments, \pm s.e.m. ($n = 5-9$). Firing rate changes were averaged over a window from 100 ms to 200 ms after valve opening. Sensitivity was not significantly different in wild type versus the *Gr28b.d* mutant (unpaired *t*-tests with iterative Bonferroni corrections, see Methods). Blocking inhibition in these cells produced oscillatory activity that precluded analysis of responses to small steps with inhibition blocked. Genotypes are: *GMR54A03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP* (wild type) and *Gr28b*^{MB03888}; *GMR54A03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP* (*Gr28b.d*^{-/-}).



Extended Data Figure 9 | Validating the *Gad1-Gal4* line as a marker of GABAergic local neurons in the posterior antennal protocerebrum. In pilot studies, we found that all the local neurons we encountered in the proximal antennal protocerebrum had somata within a distinctive cluster ventral to the antennal lobes. We therefore targeted our electrodes to this cluster when we used a *Gad1-Gal4* line²⁶ to drive CD8::GFP expression in putative GABAergic neurons. In order to determine if the CD8::GFP-expressing somata in this

cluster are indeed GABAergic, we performed dual immunofluorescence confocal microscopy with anti-CD8 and anti-GABA antibodies. This representative image shows that almost all CD8::GFP-expressing somata in this region are GABA-immunopositive (90% overall), although not all GABAergic somata express CD8::GFP. This image is a single coronal section through the cluster of somata.



Extended Data Figure 10 | Additional examples of LNs in the proximal antennal protocerebrum. For each LN, the sequence of panels is analogous to that of Fig. 4: raw traces (a, e), peri-stimulus time histograms (b, f) and sensitivity plots (c, g). The morphologies of these cells (d and h) are displayed as in Extended Data Fig. 2, with a red outline indicating the boundary of the proximal antennal protocerebrum. **a–d**, A GABAergic LN. This LN is excited by warming and is inhibited by cooling. This cell was recorded in the genotype

in *pJFRC7-20XUAS-IVS-mCD8::GFP/+;Gad1-Gal4/+*. **e–h**, A glutamatergic LN. This LN is excited by cooling and is inhibited by warming, and its properties are similar to those of all the glutamatergic LNs we recorded from. In total we recorded from six glutamatergic LNs, three in the genotype *GMR52G03-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP* and three in the genotype *GMR91H10-Gal4,pJFRC2-10XUAS-IVS-mCD8::GFP*.

Temperature representation in the *Drosophila* brain

Dominic D. Frank¹, Genevieve C. Jouandet¹, Patrick J. Kearney¹, Lindsey J. Macpherson² & Marco Gallio¹

In *Drosophila*, rapid temperature changes are detected at the periphery by dedicated receptors forming a simple sensory map for hot and cold in the brain¹. However, flies show a host of complex innate and learned responses to temperature, indicating that they are able to extract a range of information from this simple input. Here we define the anatomical and physiological repertoire for temperature representation in the *Drosophila* brain. First, we use a photolabelling strategy² to trace the connections that relay peripheral thermosensory information to higher brain centres, and show that they largely converge onto three target regions: the mushroom body, the lateral horn (both of which are well known centres for sensory processing) and the posterior lateral protocerebrum, a region we now define as a major site of thermosensory representation. Next, using *in vivo* calcium imaging³, we describe the thermosensory projection neurons selectively activated by hot or cold stimuli. Fast-adapting neurons display transient ON and OFF responses and track rapid temperature shifts remarkably well, while slow-adapting cell responses better reflect the magnitude of simple thermal changes. Unexpectedly, we also find a population of broadly tuned cells that respond to both heating and cooling, and show that they are required for normal behavioural avoidance of both hot and cold in a simple two-choice temperature preference assay. Taken together, our results uncover a coordinated ensemble of neural responses to temperature in the *Drosophila* brain, demonstrate that a broadly tuned thermal line contributes to rapid avoidance behaviour, and illustrate how stimulus quality, temporal structure, and intensity can be extracted from a simple glomerular map at a single synaptic station.

Thermosensation provides animals with critical information about their environment. We have previously shown that, in *Drosophila*, rapid temperature changes are detected by dedicated hot and cold temperature receptors in the last antennal segment, the arista, with three cells responding to warming and three to cooling¹. The projections of these sensory neurons target the base of the antennal lobe, a region previously referred to as the proximal antennal protocerebrum (recent *Drosophila* brain nomenclature⁴ refers to this region as the posterior antennal lobe). At the posterior antennal lobe, hot and cold temperature receptors form two distinct, adjacent glomeruli defining a simple map for temperature coding¹. Here we focus on identifying the second-order thermosensory neurons (referred to as thermosensory projection neurons) as a first step towards understanding how this spatially segregated pattern of activity is processed to mediate appropriate responses to temperature stimuli.

Thermosensory projection neurons (tPNs) should display prominent dendritic posterior antennal lobe (PAL; see Methods) innervation. Therefore, we reasoned that two-photon guided conversion of photoactivatable GFP (PA-GFP^{2,5}) could provide a suitable approach to identify their projections. We engineered flies where PA-GFP was constitutively expressed in all brain neurons, and the hot and/or cold temperature receptor terminals were selectively labelled by a red fluorescent protein⁶ (Extended Data Fig. 1). Using the red fluorescence as a guide, we then photo-converted PA-GFP in either the hot (Fig. 1a) or the cold (Fig. 1b) glomerulus, resulting in labelling of ascending axons. Projection neurons originating in the PAL follow several widely different trajectories to reach their targets (Fig. 1). Among these, one prominent pathway targeted the calyx (the input region of the mushroom body) and the edge of the

lateral horn, from where a subset of axons further extended ventrally to terminate in the posterior lateral protocerebrum (PLP, Fig. 1a, b, top arrowhead). Another group of tPNs initially followed the same ascending tract, but bypassed the calyx, instead veering off laterally to innervate the same domains of the lateral horn and PLP (Fig. 1a, b, centre arrowhead). A third class of tPNs directly extended axons laterally, following either an anterior or posterior path that still terminated in the same domains of the lateral horn and PLP (Fig. 1a, b, bottom arrowhead). Thus, despite following different routes, these pathways largely converged onto just three regions: the mushroom body calyx, the lateral horn (well-known centres for sensory processing in the fly brain) and the PLP, a poorly characterized region that now emerges as a major site of thermosensory signal processing. Our data also demonstrated significant overlap between projections innervating the hot and cold glomeruli. Indeed, we only observed a single modality-specific fibre that connected the cold sensory terminals to a 'microglomerulus' at the edge of the calyx (Fig. 1b, open arrowhead; this unusual fibre has been noted previously in anatomical studies⁷).

These findings raise the question of why thermosensory information is split into different pathways, which largely converge onto the same target regions. To address this question, we searched the large collections of Gal4 driver lines made available by the Janelia Farm FlyLight initiative⁷, as well as the Vienna Tile project (<http://stockcenter.vdrc.at/>), and selected lines to target specific tPN populations for functional

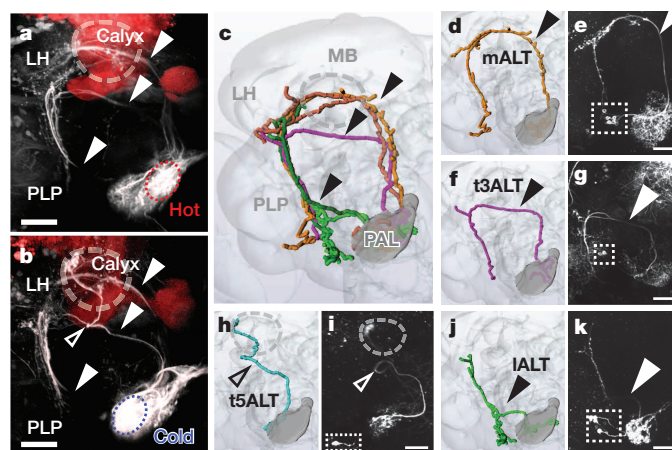


Figure 1 | Projection neuron pathways in the fly thermosensory system. a, b, PA-GFP photoconversion directed to the pre-synaptic termini of hot (a, red circle) or cold (b, blue circle) temperature receptors labels post-synaptic pathways, revealing common projections (filled arrowheads in a and b) and a single, cold-specific fibre (empty arrowhead in b), as well as a high degree of convergence on three target regions: the mushroom body (MB), lateral horn (LH) and posterior lateral protocerebrum (PLP). Mushroom bodies were labelled with dsRed as a landmark. c, d, f, h, j, Three-dimensional reconstruction of specific pathways. e, g, i, k, Two-photon stacks of brains expressing GFP under the control of drivers selectively expressed in corresponding tPNs (see Methods for genotypes). Drivers: e, R22C06; g, R84E08; i, R60H12; k, R95C02 (cell bodies are boxed). ALT, antennal lobe tract; l, lateral; m, medial; t, transverse. Scale bars, 25 μm.

¹Department of Neurobiology, Northwestern University, Evanston, Illinois 60208, USA. ²Departments of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032, USA.

studies. Visual inspection of nearly 12,000 confocal stacks (representing more than 10,000 lines) confirmed that our PA-GFP experiments had identified the major projections originating from the PAL and provided an initial list of ~200 drivers targeting tPNs. For this work, we used ten representative drivers fulfilling three important criteria: (1) selective Gal4 expression, as this would help to identify unambiguously the neurons of interest; (2) confirmed synaptic connectivity with the sensory neurons, as detected by GFP reconstitution across synaptic partners (GRASP^{8,9}, including a new synaptic variant we developed; Extended Data Fig. 2); (3) rapid functional responses to naturalistic temperature stimuli, recorded *in vivo* by gently exposing brain tissue through a hole in the head cuticle and imaging the activity of tPNs by two-photon microscopy¹. Finally, we used acute resection of the antennal nerve as a means to confirm that the cell's responses were in fact driven by the antennal temperature receptors.

The drivers listed in Extended Data Table 1 fulfil all these criteria and provide a comprehensive repertoire of thermosensory projection neurons, while the anatomy of a representative set of tPN cell types (reconstructed by transgenic labelling with GFP) is shown in Fig. 1c–k. Finally, we confirmed that all identified tPNs displayed the expected polarity of a projection neuron (that is, dendrites in the PAL and axon terminals in higher brain centres) by targeting expression of a dendritic marker (DenMark¹⁰, Extended Data Table 1) and of a pre-synaptic GFP fusion (synt-GFP¹¹, Extended Data Fig. 3). In all, our screen identified seven tPN cell types with distinct innervation patterns and functional properties (see below).

Thermoreceptor neurons in the antenna respond either to cooling or heating and define 'labelled lines' for temperature coding at the periphery¹. Functional imaging studies revealed second-order neurons that were also selectively activated by either cooling or heating (that is, 'narrowly tuned'), and specifically connected to either the cold or hot temperature receptors (as demonstrated by GRASP, Extended Data Fig. 2 and Extended Data Table 1). For example, robust, sensitive responses to cooling were reliably observed from neurons innervating the cold-specific fifth transverse antennal lobe tract (t5ALT) (Fig. 2) and displaying selective GRASP with cold temperature receptors (Extended Data Fig. 2, R60H12), while we recorded robust heating responses from cells innervating the lateral ALT pathway and selectively connecting with hot temperature receptors (VT46265; a full description of the properties of the various cell types is provided in Extended Data Table 1).

Narrowly tuned projection neurons could be categorized based on the decay profile of their calcium responses as either 'slow-adapting' or 'fast-adapting'. Slow-adapting tPNs, such as the cold-specific t5ALT tPN, responded to temperature stimuli with calcium transients that persisted during the stimulus and even after the temperature had returned to baseline (Fig. 2b, arrowheads). As shown in Fig. 2d, the peak responses of this cell type scaled with the magnitude of cooling stimuli over a wide range of intensities. Yet, as a consequence of slow decay, intracellular calcium did not return to baseline when cooling stimuli were rapidly interleaved (Fig. 2e).

In contrast, fast-adapting cells responded to temperature changes with a calcium transient that did not faithfully scale with stimulus intensity and which was followed by fast decay, as illustrated in Fig. 3 for a hot tPN innervating the lateral pathway (Fig. 3a–d; see Extended Data Fig. 4 for a comparison of fast- and slow-adapting cold cells). As a result of fast kinetics, the peak response of this cell type generally preceded the stimulus peak (Fig. 3d) and, for larger stimuli, intracellular calcium had nearly returned to the pre-stimulus baseline when the temperature was still rapidly changing (Fig. 3c). Because of this, these fast-adapting cells are unlikely to code information regarding the peak temperature of the stimulus (Fig. 3e), yet they were able to track remarkably well a rapidly evolving temperature transient (Fig. 3f).

One of the drivers that we identified is active in a group of six such fast-adapting neurons, four of which are activated by cooling and two by heating, allowing one to simultaneously record the responses of both cell types under two-photon microscopy. Our 'hot' stimuli consist of a

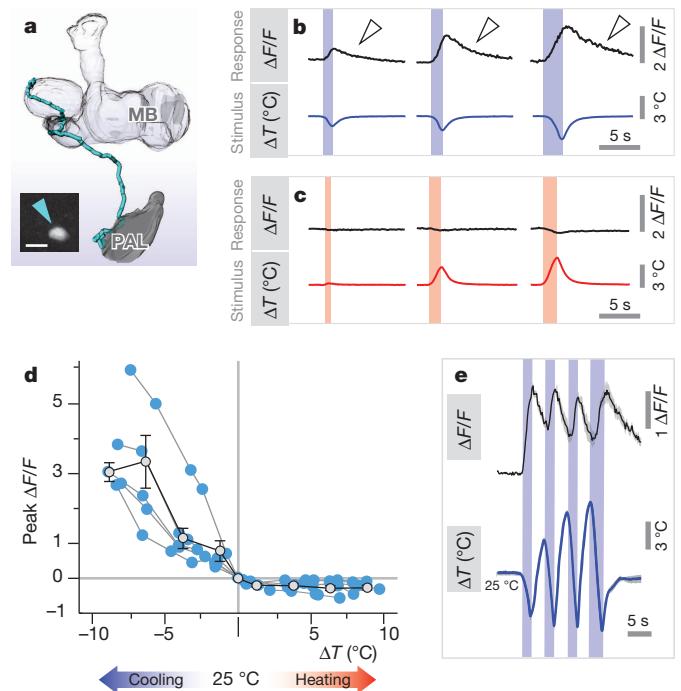


Figure 2 | Properties of slow-adapting, cold-activated projection neurons. **a**, Three-dimensional reconstruction of a tPN that innervates the t5ALT tract (expressing R60H12; cell body in inset). **b**, **c**, Representative calcium-imaging traces from this tPN. Robust transients are seen in response to cooling stimuli of increasing magnitude (**b**), but not in response to heating (**c**). Notably, intracellular calcium concentration ($[Ca^{2+}]_i$) remains above baseline well after the end of cooling (arrowheads in **b**). **d**, Stimulus–response plot. Blue dots, peak response of a single cell to a defined stimulus, connected dots belong to the same cell, $n = 5$ cells/5 animals. Grey dots, binned averages \pm s.e.m.; bin width = 2.5°C , bin at zero = 0°C ΔT . Responses are proportional to cold stimuli ($R^2 = 0.5$, $P < 0.001$; slope = -0.43 , 95% confidence interval (CI) = $[-0.61$ to $-0.25]$) but not to hot stimuli (R^2 P = not significant). **e**, $[Ca^{2+}]_i$ does not return to baseline when cooling stimuli are rapidly interleaved with heating ones (black trace = average of 3 responses, 3 cells/3 animals \pm s.e.m., in grey; blue trace = average of the 3 stimuli \pm s.e.m., grey). In all panels $\Delta F/F$ indicates fluorescence change over baseline; ΔT indicates temperature change. Scale bar in **a**, $5\ \mu\text{m}$.

heating pulse followed by cooling, which quickly brings the temperature back to baseline. As expected, we observed a transient calcium response in the hot-activated cell type at the beginning of the heating step (Fig. 3g–i, 'ON' response). Interestingly, the cold-activated cell type did not immediately respond at the onset of the following cooling phase (as would be expected for a simple cooling response), but rather with a significant delay; that is, at the very end of the temperature transient when the temperature was again approaching baseline ('OFF' response, Fig. 3i). Even in the midst of a rapidly varying temperature transient, ON and OFF cell responses alternated, with hot cells tracking the onset of heating, and cold cells becoming active when the temperature approached the 25°C baseline (Extended Data Fig. 5).

We note that the specific delay of OFF responses was observed only when cooling was preceded by heating, perhaps reflecting inhibitory interactions between hot and cold processing circuits. Furthermore, OFF responses coincided with the return to baseline even when this was systematically varied between $\sim 20^\circ\text{C}$ and $\sim 30^\circ\text{C}$ (Extended Data Fig. 6). Hence, OFF responses appear to demarcate return to baseline after a rapid change, and might for example help the animal maintain course on a defined thermocline. Hot-activated projection neurons also displayed OFF responses during the heating phase that followed a cooling stimulus, but the amount of delay varied between cells (Extended Data Fig. 6).

Taken together, our results suggest that the distinct kinetics of slow- and fast-adapting tPNs may tune each class to relay significantly different

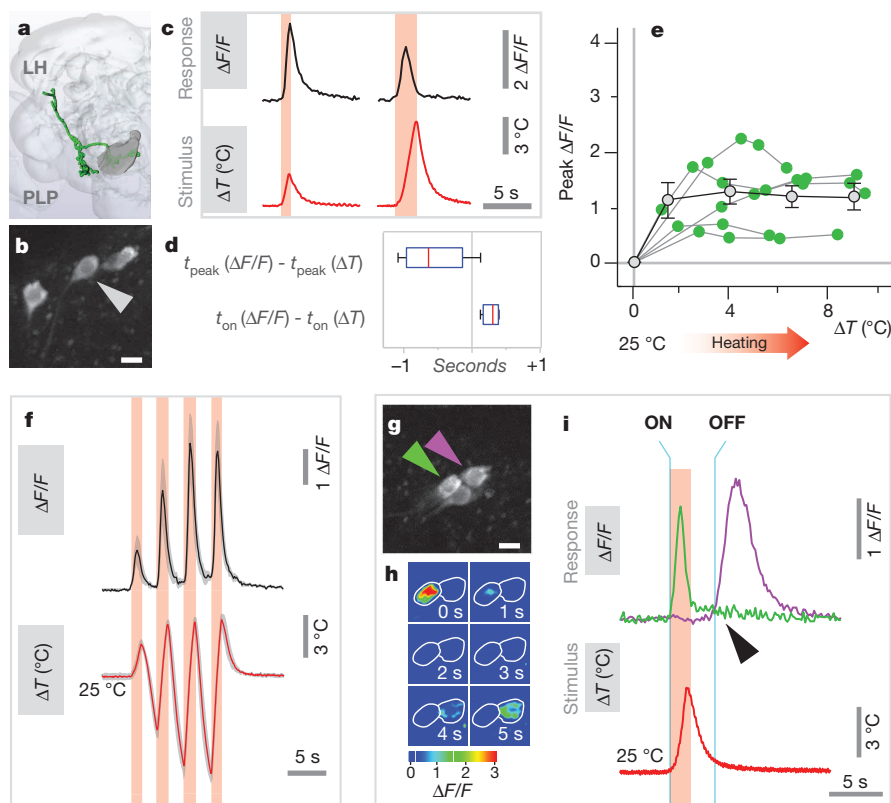


Figure 3 | Fast-adapting projection neurons display ON and OFF responses to temperature stimuli. **a**, Fast-adapting tPNs responding to either hot or cold innervate a lateral pathway (three-dimensional reconstruction in green). **b**, **c**, A single cell in this cluster (expressing R95C02) displays rapid but transient responses to hot stimuli of different amplitude. Note that, for the larger stimulus, the response has already nearly returned to baseline while the temperature is still changing. **d**, Box-plot showing the temporal difference between response and stimulus peak (t_{peak}) and onset (t_{on}), respectively. Owing to its transient nature, the response often peaks before the stimulus (red line, median; blue box, 25th and 75th percentiles; whiskers, data range; $n = 15$ stimuli, $\Delta T = 3.4 \pm 0.9$, mean \pm s.d.; 5 cells/5 animals). **e**, Stimulus-response plot for defined heating stimuli (green dots, peak response of a single cell to a stimulus, connected dots belong to the same cell, $n = 5$ cells/5 animals; grey dots, binned averages \pm s.e.m.; bin width = 2.5°C , bin at zero = 0°C ΔT). Responses of this cell type do not correlate with stimulus intensity ($R^2 P =$ not significant). **f**, This cell type tracks well a dynamic temperature stimulus (black trace, average of 3 responses, 3 cells/3 animals \pm s.e.m., in grey; red trace, average of the 3 stimuli \pm s.e.m., grey). **g–i**, The driver R95C02 is expressed in both hot- (green arrow in **g** and traces in **i**) and cold-activated tPNs (purple arrow in **g** and traces in **i**). **h**, A sequence of $\Delta F/F$ images of two cell bodies responding to a hot stimulus ($\Delta T \sim 5^\circ\text{C}$, peaking at 0 s, same cells as in **g**); **i**, single stimulus and response traces from the two cell bodies in **g**, recorded simultaneously. **i**, The hot activated tPN responds to a heating stimulus with a rapid ON transient, while the cold cell displays a delayed OFF response. Scale bars in **b** and **g**, $5\ \mu\text{m}$.

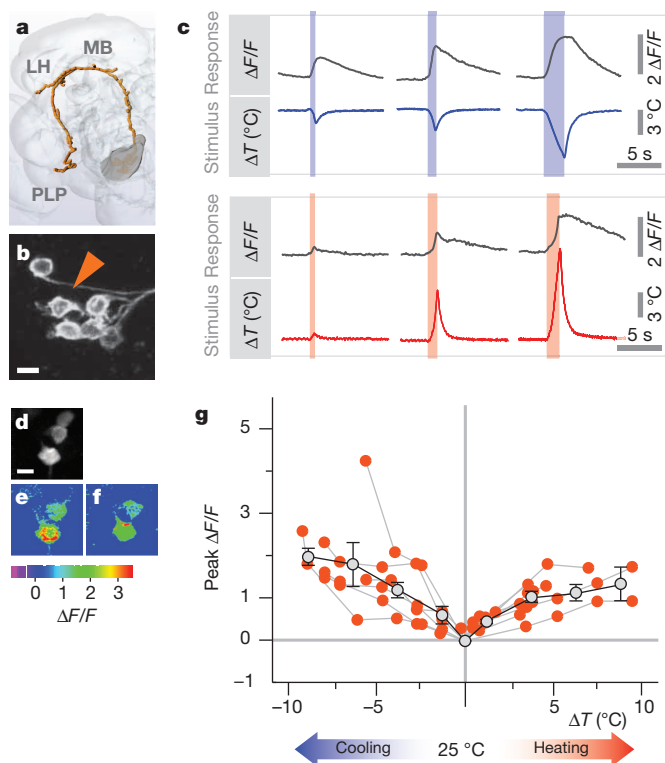


Figure 4 | Broadly tuned thermosensory projection neurons respond to both heating and cooling stimuli. **a**, Three-dimensional reconstruction of a broadly tuned tPN pathway (VT40053; cell bodies in **b**). **c**, This cell type responds with transients of increasing magnitude to both cooling (top) and heating stimuli (bottom) of increasing intensity (shown are representative traces). **d–f**, A maximum-response image clearly shows that the same neurons are capable of both cooling (**e**) and heating (**f**) responses ($\Delta T \approx \pm 5^\circ\text{C}$). **g**, Stimulus-response plot for defined heating and cooling stimuli (orange dots, peak response of a single cell to a stimulus, connected dots belong to the same cell, $n = 6$ cells/5 animals; grey dots, binned averages \pm s.e.m.; bin width = 2.5°C , bin at zero = 0°C ΔT); the responses are proportional to the stimuli in both the cold and hot range (cold, $R^2 = 0.34$, $P < 0.001$, slope = -0.20 , 95% CI = $[-0.31 \text{ to } -0.09]$; hot, $R^2 = 0.5$, $P < 0.001$, slope = 0.12 , 95% CI = $[0.07 \text{ to } 0.17]$). Scale bars in **b** and **d**, $5\ \mu\text{m}$.

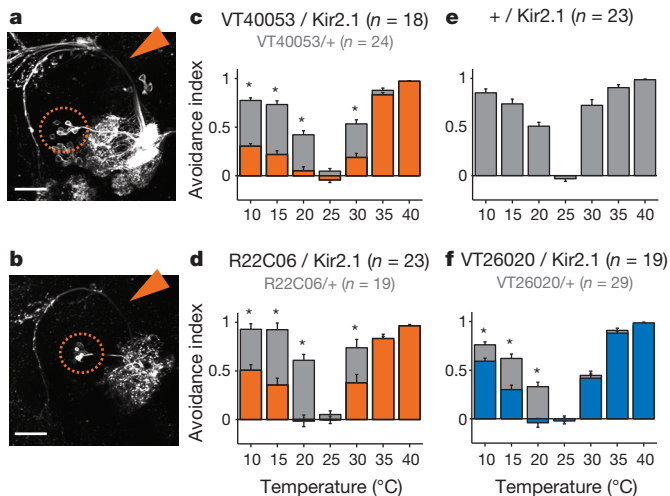


Figure 5 | Both broadly and narrowly tuned tPNs are required for rapid temperature avoidance in a two-choice behavioural assay. **a, b**, The VT40053 (**a**) and R22C06 (**b**) drivers are prominently expressed in broadly tuned tPNs, as shown by GFP expression (cell bodies are circled; arrows indicate mALT; scale bars, 25 μ m). **c, d**, Silencing this cell type produces flies with significant defects in both hot and cold avoidance in a two-choice assay. **e**, Avoidance indices of control flies bearing the Kir2.1 construct but no driver. **f**, Silencing three distinct cold-activated tPNs (expressing VT26020) results in a selective defect in cold avoidance (scores of the drivers alone are shown in **c, d** and **f** as grey bars; avoidance index values were compared by two-way ANOVA, asterisks denote a significant interaction between the Gal4 and UAS transgenes, $P < 0.01$; error bars indicate mean \pm s.e.m.; the number of biological replicates is in parentheses next to each genotype).

aspects of thermal stimuli to their targets: slow-adapting neurons appear better suited to convey the magnitude of a temperature change, while fast-adapting cells may better track its temporal structure (onset and offset).

In addition to ‘narrowly tuned’ cells (that is, responding to either hot or cold), we also discovered a significant class of ‘broadly tuned’ tPNs (Fig. 4). Each neuron in this group displayed calcium transients in response to both heating and cooling, the peak of which correlated well with the stimulus intensity in either direction (Fig. 4c, g). Unexpectedly, broadly tuned tPNs represented the largest group of cells we found (~10 cells, a significant number when compared to other projection neurons; see Extended Data Table 1) and innervated all major targets we described (mushroom body, lateral horn and PLP). Therefore, these cells are likely to provide significant drive to higher-order thermosensory regions during any temperature transient.

We have previously shown that rapid temperature preference is mediated by the opposing pushes of heat- and cold-aversion¹. One attractive hypothesis is that these neurons could represent a common aversive behavioural line. If this is the case, silencing their activity should dampen aversive responses to both hot and cold stimuli. To test this, we selected two drivers prominently expressed in these cells for functional studies: VT40053 is active in all the broadly tuned tPNs, but its expression also includes additional sensory circuits (Fig. 5a). R22C06 is only expressed in about half of the broadly tuned cells, but in very few other cells in the brain or ventral nerve cord (Fig. 5b, and see Methods for a detailed description of these lines). Importantly, with the exception of tPNs, these two drivers show no overlap in expression.

Driving Kir2.1 (ref. 12) (that is, silencing neural activity by hyperpolarization) under the control of either line resulted in a significant decrease in the avoidance of both cold and hot regimes in a rapid two-choice test¹ (Fig. 5c, d), demonstrating that this pathway is indeed required for rapid temperature preference. In contrast, expression of Kir2.1 under the control of a driver active in three distinct cold-activated tPNs

also innervating the medial ALT pathway (VT26020, see Extended Data Fig. 7 and Methods for details) produced an avoidance defect that was highly selective for cold temperatures (Fig. 5f), suggesting that both narrowly and broadly tuned tPNs participate in temperature preference. We note that flies in which broadly tuned tPNs were inactivated still avoided hot temperatures at the upper limit of their tolerance range ($>35^{\circ}\text{C}$, Fig. 5), indicating that their locomotor/navigational programs are intact, and that the avoidance of high temperatures probably involves additional circuit components.

We define here a circuit diagram for the representation of temperature in the *Drosophila* brain (Extended Data Fig. 8). ‘Hubs’ of this circuit are the PAL, where the activity of hot and cold glomeruli initially signals a temperature change, and the triad of mushroom body, lateral horn and PLP, where thermosensory stimuli are richly represented by an array of differently tuned projection neurons. Taken together, our results demonstrate that substantial processing of thermosensory input at a single synaptic step can extract information about the sign, onset, magnitude and duration of a temperature change, and establish a framework to understand how complex temperature responses, often happening on different timescales, can be orchestrated starting from a simple sensory map.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 July 2014; accepted 5 February 2015.

Published online 4 March 2015.

- Gallio, M., Ofstad, T. A., Macpherson, L. J., Wang, J. W. & Zuker, C. S. The coding of temperature in the *Drosophila* brain. *Cell* **144**, 614–624 (2011).
- Ruta, V. *et al.* A dimorphic pheromone circuit in *Drosophila* from sensory input to descending output. *Nature* **468**, 686–690 (2010).
- Wang, J. W., Wong, A. M., Flores, J., Vosshall, L. B. & Axel, R. Two-photon calcium imaging reveals an odor-evoked map of activity in the fly brain. *Cell* **112**, 271–282 (2003).
- Ito, K. *et al.* A systematic nomenclature for the insect brain. *Neuron* **81**, 755–765 (2014).
- Datta, S. R. *et al.* The *Drosophila* pheromone cVA activates a sexually dimorphic neural circuit. *Nature* **452**, 473–477 (2008).
- Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnol.* **22**, 1567–1572 (2004).
- Jenett, A. *et al.* A GAL4-driver line resource for *Drosophila* neurobiology. *Cell Rep.* **2**, 991–1001 (2012).
- Feinberg, E. H. *et al.* GFP Reconstitution Across Synaptic Partners (GRASP) defines cell contacts and synapses in living nervous systems. *Neuron* **57**, 353–363 (2008).
- Gordon, M. D. & Scott, K. Motor control in a *Drosophila* taste circuit. *Neuron* **61**, 373–384 (2009).
- Nicolai, L. J. *et al.* Genetically encoded dendritic marker sheds light on neuronal connectivity in *Drosophila*. *Proc. Natl Acad. Sci. USA* **107**, 20553–20558 (2010).
- Zhang, Y. Q., Rodesch, C. K. & Broadie, K. Living synaptic vesicle marker: synaptotagmin-GFP. *Genesis* **34**, 142–145 (2002).
- Baines, R. A., Uhler, J. P., Thompson, A., Sweeney, S. T. & Bate, M. Altered electrical properties in *Drosophila* neurons developing without synaptic transmission. *J. Neurosci.* **21**, 1523–1531 (2001).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank V. Ruta and B. Noro for providing PA-GFP flies and reagents, and especially G. Rubin for providing access to the FlyLight collection before publication. We are grateful to N. Ryba, D. Yarmolinsky, C. Zuker and members of the Gallio laboratory for critical comments on the manuscript. Z. Turan and A. Kuang provided technical assistance and we thank C. Ulmer and a number of undergraduate students for fly husbandry. This work was supported by NIH grant 1R01NS086859-01 (to M.G.) and by training grant 2T32MH067564 (to D.D.F.).

Author Contributions M.G. and D.D.F. designed the study, carried out the imaging experiments, analysed data (with help from G.C.J.), and wrote the paper; D.D.F., G.C.J. and M.G. ran and analysed all behavioural experiments; P.J.K. carried out GRASP experiments using transgenic lines produced by M.G. and L.J.M.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to M.G. (marco.gallio@northwestern.edu).

METHODS

No statistical methods were used to predetermine sample size.

Experimental animals and transgenes. The following transgenic flies were constructed for this study: *GR28b.d-LexA*, expressing LexA under a promoter element active in hot receptor neurons^{13,14}; *IR94a-LexA*, from a promoter expressed in both hot and cold cells (additionally expressed in a small number of olfactory sensory neurons¹⁵). *Aop-syb:spGFP 1-10*, a split-GFP 1-10 fragment targeted to pre-synaptic sites by fusion to the cytoplasmic domain of synaptobrevin, for GRASP (ref. 8; L.J.M. and M.G., unpublished; see also ref. 16). Additional transgenic lines: *Aop-CD4:spGFP 1-10*, *UAS-CD4:spGFP 11* (ref. 9) (a gift from K. Scott); *UAS-SPA* (ref. 2); *UAS-C3PA* (ref. 2), *MB-DsRed*, *syb-Gal4* (all a gift from B. Noro and V. Ruta); *UAS-Kir2.1* (ref. 12), *UAS-CD8:GFP*, *Aop-Tdtomato* (ref. 6), *Aop-CD2:GFP*, *UAS-GCaMP6m* (ref. 17), *UAS-DenMark* (ref. 10), *UAS-syt:GFP* (ref. 11). Gal4 driver lines were obtained from the Bloomington Stock centre or VDRC. A full list of genotypes is provided as Supplementary Information.

Two-photon guided conversion of photo-activatable GFP in PAL glomeruli.

To visualize axonal projections arising from the PAL, we used two-photon guided conversion of PA-GFP following the protocol of Ruta *et al.*². All imaging was conducted on a Prairie Ultima two-photon microscopy system equipped with a Coherent Chameleon Ti:Sapphire laser, GaAsP PMT, and an Olympus 40× 0.9NA or 60× 1.1NA water immersion objective. In brief, we labelled hot temperature receptor terminals alone or both hot and cold ones by expression of fluorescent proteins (*tdTomato* (ref. 6) or GFP; under LexA, see above). We dissected 2–4-day-old adult flies, and defined 3D ROIs around selected glomeruli while imaging at 945 nm (a wavelength that does not activate PA-GFP). Finally, the laser was tuned to 720 nm and the volume scanned at low power (10–30 mW, measured at the back aperture of the objective) 30 times interleaved by 30 s wait periods. A ~15-min rest allowed diffusion of photo-converted PA-GFP.

Nomenclature of brain regions and 3D reconstruction of axonal pathways. All efforts were taken to adhere to the standardized insect brain nomenclature described in Ito *et al.*⁴. The area defined by the glomeruli formed by hot and cold sensory neurons (previously referred to as the PAP¹) is here referred to as the posterior antennal lobe (PAL), as the two regions have been now merged. Indeed, the projections of arista neurons were first described by Stocker *et al.* as targeting the posterior AL¹⁸ (glomeruli VP2 and VP3). Interestingly, our analysis revealed an additional glomerulus close to the PAL (probably VP1) that seems to be innervated by cold-responding sacculus neurons (data not shown). However, it was also activated by non-thermal stimuli and therefore excluded from further analysis. The mAL, mlAL and lAL fibre tracts are described in Ito *et al.*⁴. The transverse 3AL and 5AL tracts have been described in Tanaka *et al.*¹⁹ and Jenett *et al.*⁷, respectively. 3D reconstructions of axonal tracts (Fig. 1) were carried out on 3D brain volumes pre-aligned to a standard brain, available through BrainBase²⁰. Semi-automated neurite reconstruction was carried out by AMIRA on representative ViennaTile driver lines, and information from PA-GFP experiments and additional driver lines used to refine neurite models by hand. Automated annotation of the projections of driver lines in FlyLight and/or BrainBase directly suggested innervation of target regions (mushroom body, lateral horn and PLP) based on volume reconstructions. We adopted this nomenclature (even when, for example in the case of the lateral horn, tPN innervation appears confined to the lateral horn margin), and confirmed the presence of putative pre-synaptic terminals by expression of syt-GFP (Extended Data Fig. 3). 3D reconstructions of the mushroom body and lateral horn shown in Extended Data Fig. 3 were carried out in AMIRA based on the nc82/bruchpilot staining of the standard brain; because the boundaries of the PLP could not be easily defined in nc82/bruchpilot staining, the PLP was not reconstructed but is shown as an outline.

GRASP, Denmark and syt-GFP visualization and staining. Native GRASP fluorescence was visualized at the synapses between hot, and hot and cold receptor neurons and their potential projection neurons in the PAL using two-photon microscopy of 2–4-day-old adult dissected fly brains. Immunofluorescence of GRASP was performed on 4% PFA-fixed brains using antibodies which specifically either bind to reconstituted GRASP⁹–mouse anti-GFP (1:100; Sigma, catalogue no. G6539), or that preferentially recognizes spGFP10–chicken anti-GFP (1:1000, abcam 13970). To determine the polarity of tPNs and identify putative pre-synaptic termini, we expressed under the control of tPN Gal4 lines either the red fluorescent dendritic marker DenMark¹⁰ or a synaptotagmin–GFP¹¹ fusion, respectively, and visualized native fluorescence by confocal (DenMark) or 2-photon (syt:GFP) microscopy.

Functional calcium imaging of tPN responses to temperature stimuli. Our calcium imaging and temperature stimulation set-up has been previously described in detail¹. In brief, UAS.GCaMP6m flies or intact fly heads were submerged in AHL³ and sufficient cuticle surrounding the brain was delicately removed to provide optical access to the projection neurons. The preparation was then placed in a custom-built microfluidics chamber, covered with a plastic coverslip, and placed on the two-photon microscope stage. Rapid temperature changes were achieved by controlling

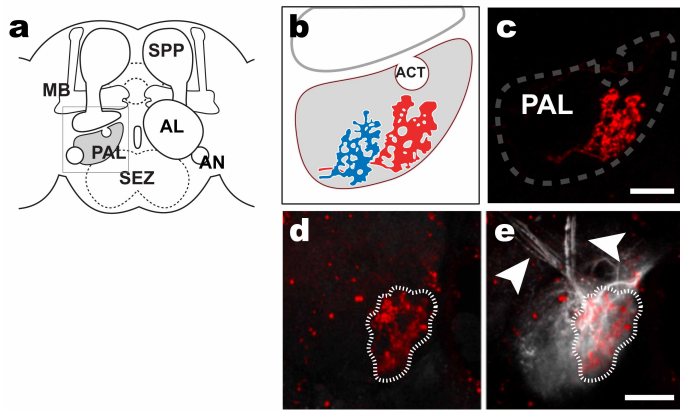
the temperature of the medium via a custom-built system consisting of a series of three-way valves (Lee Instruments, response time 2 ms) and Peltier elements independently controlling baseline, 'hot' and 'cold' flow, respectively. 'Complex' stimuli (Figs 2 and 3 and Extended Data Fig. 5) were obtained by actuating the control valves using a sequence of TTL pulses to alternate between hot flow and cold flow. The delay between valve trigger and onset of ΔT was a function of tubing length, 1–2 s. Baseline experiments (Extended Data Fig. 6) were done by adjusting the temperature of baseline flow and allowing the preparation to adapt to it for ~1 min before recordings. We used rapid heating and cooling pulses¹ (~2 °C per second), as this regime closely approximates what the fly might encounter during a navigational saccade. Temperature was recorded using a BAT-12 electronic thermometer (time constant 0.004 s, accuracy 0.01 °C, Physitemp). Images were acquired at 256 × 256 pixel resolution at a rate of 4 Hz on a Prairie Ultima 2-photon microscope equipped with a Coherent Chameleon Ti:Sapphire laser, GaAsP PMT, and an Olympus 40× 0.9NA or 60× 1.1NA water immersion objective. $\Delta F/F$ analysis was carried out using custom scripts in MATLAB, base fluorescence was calculated using all frames preceding valve trigger (occurring at 5 s). All recordings were from cell bodies unless otherwise specified. To extract the 'onset' of a given response (Fig. 3e and Extended Data Fig. 5b, c) we wrote a script that identifies the first sequence of three data points in the trace all of which are two standard deviations above average background ('noise', calculated from all data points preceding valve trigger); the first point in this sequence was used as 'onset' and mapped on the corresponding time stamp. Temperature was recorded at the fly head during each experiment, and was synchronized with imaging by starting the temperature recording and image acquisition by the same trigger signal. As the recording of the two variables was synchronous (and the distance of the thermo probe from the fly head small and relatively constant), we generally observed good time-locking between response and stimulus onset. To calculate decay constants of 'fast' and 'slow' responses (τ) we confirmed that the cellular response data fit an exponential decay function, and measured the time from peak to 1/e (~37%) of peak value.

Dose–response statistics. Each point in the dose–response plots (Figs 2d, 3e and 4g) represents the peak response of one cell body to one presentation of a given temperature change, as in the corresponding representative traces shown (Figs 2b, 3c and 4c). The same cell is challenged with a set of 5–10 stimuli and the resulting responses are connected by a line ($n = 5$ cells/5 animals). For each cell type, we confirmed that the effect of temperature was significant as follows: the bin at zero contains the 'responses' recorded in control experiments ($[Ca^{2+}]_i$ maxima of traces acquired to control for valve artefacts; that is, triggering the valve system without feeding power to the Peltier elements; $\Delta T = 0$). We performed two sample *t*-tests comparing the zero bin against binned responses to temperature stimuli of ~0.5–3 °C (for all cell types, $P < 0.05$). We then calculated regression and slope statistics on all data points across individual flies to test whether the responses were proportional to stimulus intensity (see figure legends).

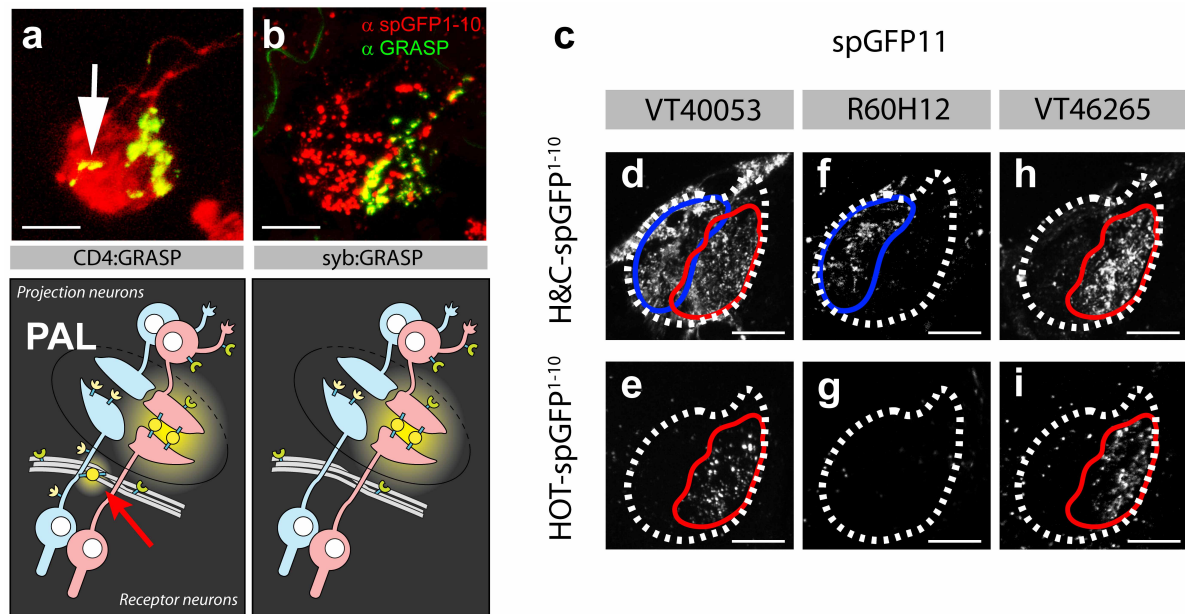
Two-choice behavioural assays for temperature preference. The two-choice assay for temperature preference has been described before¹. Briefly, 20 3–5-day-old flies were ice-anaesthetized and placed in an arena consisting of four 1' square, individually addressable Peltier tiles. In each trial, flies were presented for 3 min with a choice between 25 °C and a test temperature between 10 °C and 40 °C at 5 °C intervals, and the position of flies was recorded during each trial (by a BASLER A601FM camera) to calculate an avoidance index for each test temperature ($AI = (\text{no. flies at } 25^\circ\text{C} - \text{no. flies at test temp}) / \text{total no. flies}$). Avoidance index values for each genotype were compared by two-way ANOVA, and asterisks in Fig. 5 denote a statistically significant interaction between the Gal4 and UAS transgenes (threshold $P = 0.01$, $n > 18$ biological replicates, see figure for specific n values). Kolmogorov–Smirnov tests were used to confirm a normally distributed sample. Homogeneity of variance for each data set was confirmed by calculating the Spearman correlation (ρ) between the absolute values of the residual errors and the observed values of the dependent variable (threshold $P = 0.05$). Statistical analysis was carried out in MATLAB; sample sizes were chosen to reliably measure experimental parameters. Experiments did not involve randomization or blinding. The driver VT40053 is expressed in nearly all the broadly tuned tPNs that we have observed (~10) but its expression also includes one cold-activated tPN, olfactory projection neurons¹⁹, fibres targeting auditory centres²¹, as well as sensory projections from the taste system²² (see <http://brainbase.imp.ac.at/>). In contrast, R22C06 is selectively expressed in a small group of projection neurons projecting to the antennal lobe: a subset of the broadly tuned tPNs (5–6), 1–2 cold-activated tPNs and 1–2 olfactory projection neurons, but no other cell in the brain (and only two neurons in the ventral nerve cord). VT26020 is expressed in three cold-activated tPNs innervating the mALT whose axons do not descend to the PLP (see Extended Data Figs 2d, e and 6). Expression is also prominent in taste and auditory centres.

13. Thorne, N. & Amrein, H. Atypical expression of *Drosophila* gustatory receptor genes in sensory and central neurons. *J. Comp. Neurol.* **506**, 548–568 (2008).

14. Ni, L. *et al.* A gustatory receptor paralogue controls rapid warmth avoidance in *Drosophila*. *Nature* **500**, 580–584 (2013).
15. Benton, R., Vannice, K. S., Gomez-Diaz, C. & Vosshall, L. B. Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136**, 149–162 (2009).
16. Karuppururai, T. *et al.* A hard-wired glutamatergic circuit pools and relays UV signals to mediate spectral preference in *Drosophila*. *Neuron* **81**, 603–615 (2014).
17. Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295–300 (2013).
18. Stocker, R. F., Lienhard, M. C., Borst, A. & Fischbach, K. F. Neuronal architecture of the antennal lobe in *Drosophila melanogaster*. *Cell Tissue Res.* **262**, 9–34 (1990).
19. Tanaka, N. K., Endo, K. & Ito, K. Organization of antennal lobe-associated neurons in adult *Drosophila melanogaster* brain. *J. Comp. Neurol.* **520**, 4067–4130 (2012).
20. Bruckner, S. *et al.* BrainGazer—visual queries for neurobiology research. *IEEE Trans. Vis. Comput. Graph.* **15**, 1497–1504 (2009).
21. Lai, J. S., Lo, S. J., Dickson, B. J. & Chiang, A. S. Auditory circuit in the *Drosophila* brain. *Proc. Natl Acad. Sci. USA* **109**, 2607–2612 (2012).
22. Wang, Z., Singhvi, A., Kong, P. & Scott, K. Taste representations in the *Drosophila* brain. *Cell* **117**, 981–991 (2004).

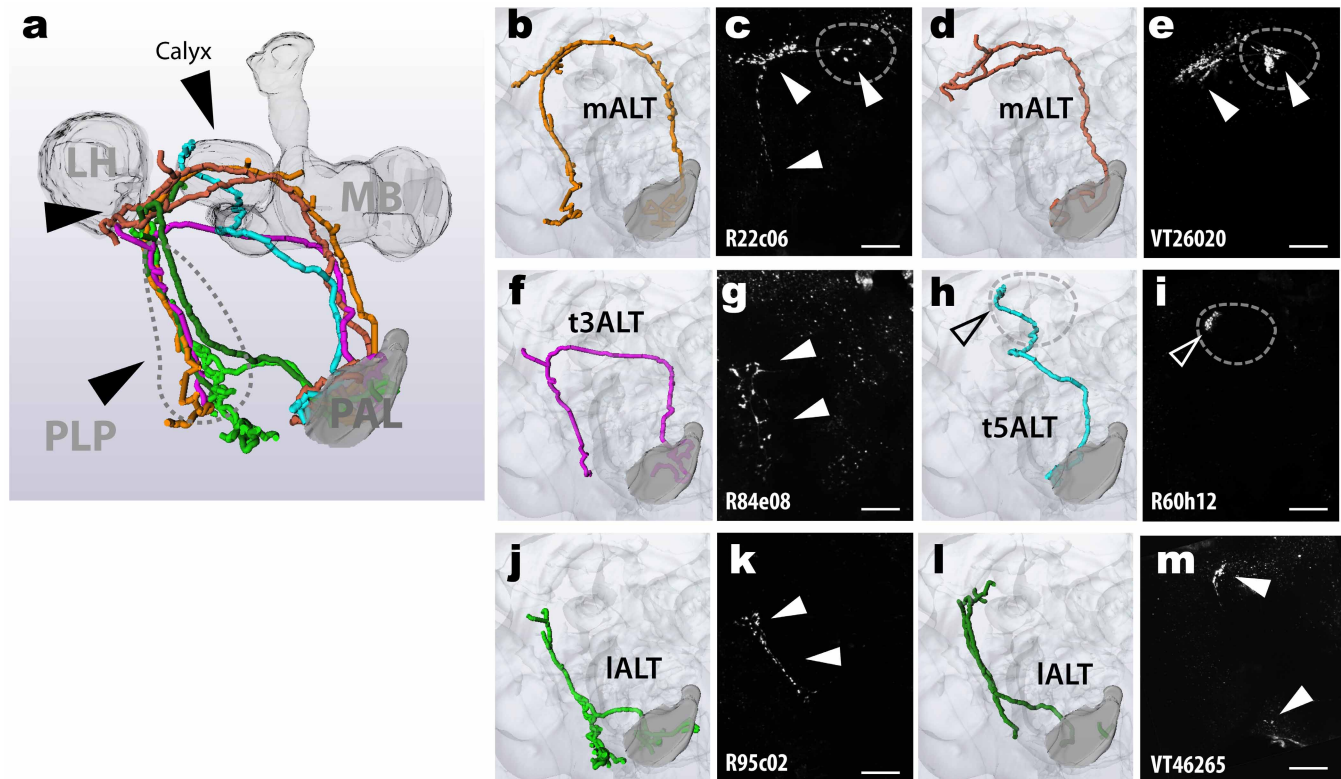


Extended Data Figure 1 | Targeted photo-conversion of PA-GFP reveals candidate tPN pathways. **a**, Diagram of the fly brain showing the location of the PAL. **b**, Schematic diagram showing the expected location of 'cold' and 'hot' glomeruli (blue and red, respectively). **c–e**, The hot glomerulus is visualized by tdTomato, driven by Gr28b.d-LexA (red). **e**, Targeted photo-activation of pan-neuronal PA-GFP (syb-Gal4/UAS-SPA) reveals ascending pathways (white arrowheads). Panels **d** and **e** are imaged under the same conditions, before (**d**) and after (**e**) photoconversion. MB, mushroom body; PAL, posterior antennal lobe; SEZ, sub oesophageal zone; AL, antennal lobe; AN, antennal nerve; SPP, super peduncular protocerebrum. Scale bars, 25 μ m.



Extended Data Figure 2 | Synaptic-targeted GFP reconstitution across synaptic partners (GRASP) reveals specific connectivity patterns between antennal temperature receptors and tPN classes. Using an n-synaptobrevin fragment to target spGFP1-10 to the pre-synaptic membrane (see Methods) increases the specificity of GRASP fluorescence at synapses between thermosensory neurons and projection neurons in the PAL. **a**, Immunofluorescence staining of GRASP contacts between hot and cold receptor neurons (IR94a-LexA, Aop-CD4:spGFP1-10) and projection neurons (VT46265, UAS-CD4:spGFP11) shows diffuse expression of spGFP1-10 throughout the membranes of the hot and cold receptor neurons (red), in addition to the GRASP signal at the synapses between hot cells and VT46265 projection neurons (green). Some non-synaptic GRASP fluorescence is seen

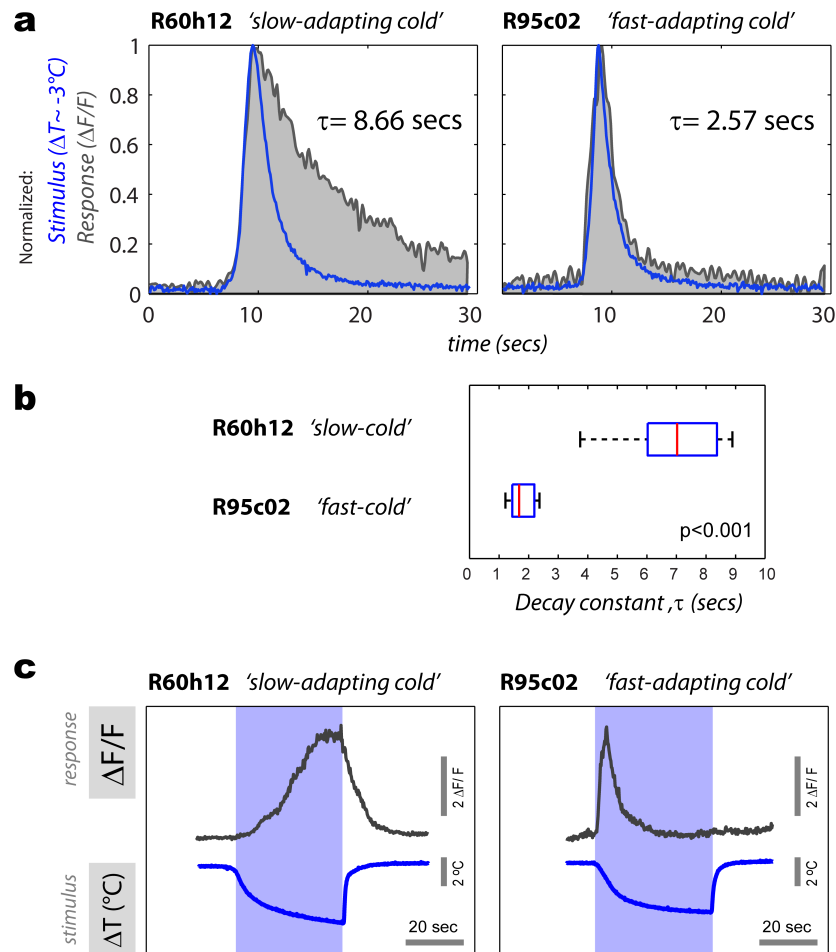
at sites of contact between the receptor neurons and fibres of passage (arrow in **a**, and see cartoon below for illustration). In contrast, synaptically targeted syb-GRASP (or syb:GRASP) (**b**) shows punctate expression of syb-spGFP1-10 (red) and specific GRASP signal (green) only at the hot-cell synapses of the PAL. **c**, Representative native GRASP fluorescence between either (**d**, **f**, **h**) hot and cold temperature receptors (IR94a-LexA; Aop-CD4:spGFP1-10) or hot only temperature receptors (**e**, **g**, **i**, GR28b.d-LexA; Aop-syb-spGFP1-10) onto three projection neuron types (VT40053, R60H12 and VT46265, each driving UAS-CD4:spGFP11). GRASP connectivity corresponds well to their tuning (see text and Extended Data Table 1 for details). Scale bars in **d–i**, 25 μ m.



Gal4> syt:GFP

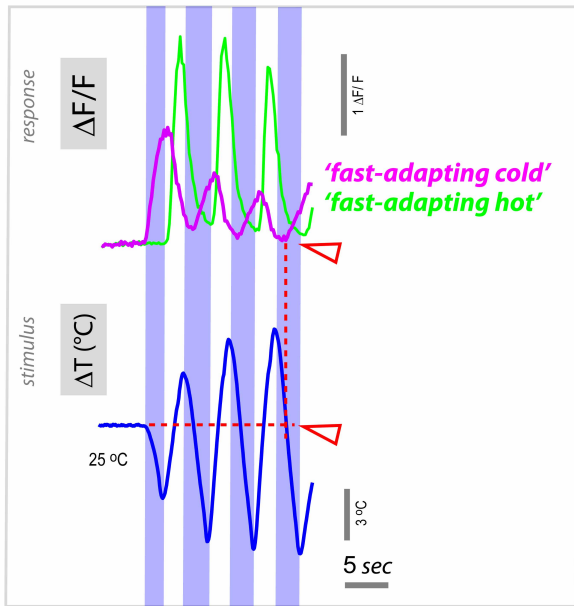
Extended Data Figure 3 | Target regions of tPN cell types. Transgenic expression of synaptotagmin–GFP reveals the location of putative pre-synaptic termini for each of the identified tPN cell types. **a**, Three-dimensional reconstruction of key target regions (the mushroom body and lateral horn, based on nc82 staining) showing their relationship to the tPN pathways (colour and abbreviations are as in Fig. 1; note that the boundaries of the PLP are not clearly demarcated by nc82 staining, hence the PLP was not

reconstructed but outlined; see Methods for details). **b, d, f, h, j, l**, Three-dimensional reconstructions of specific tPN pathways; **c, e, g, i, k, m**, two-photon stacks of brains expressing synaptotagmin–GFP under the control of Gal4 drivers expressed in neurons innervating such pathways (the driver name appears at the bottom of each panel). GFP fluorescence identifies putative pre-synaptic structures in the target regions (arrowheads). In **c, e** and **i**, the mushroom body calyx is circled. Scale bars, 25 μ m.

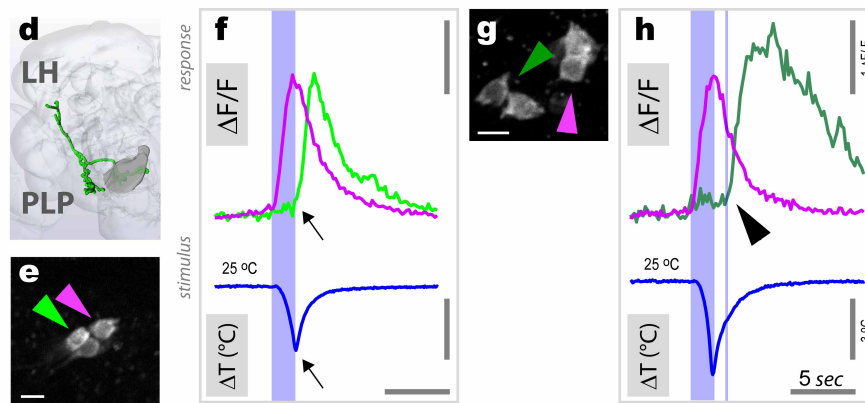
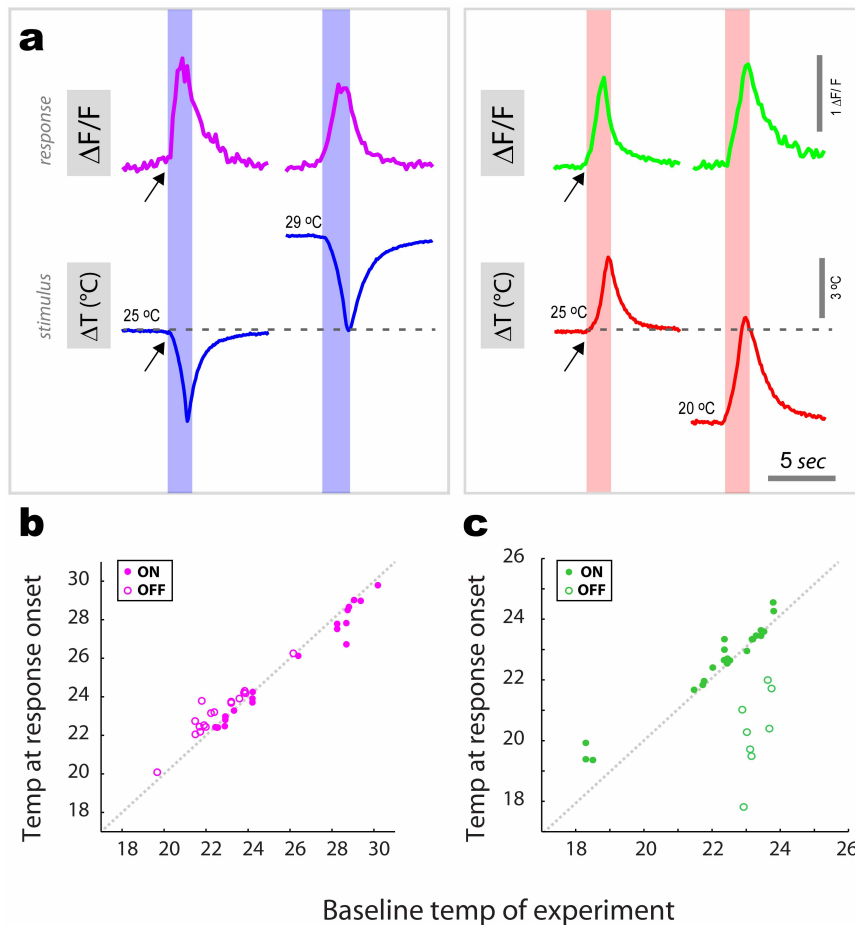


Extended Data Figure 4 | Temporal dynamics of slow- and fast-adapting cold-activated tPNs. Calcium transients of slow- and fast-adapting cold-activated tPNs differ significantly in the speed of decay. **a**, The response of a fast tPN (right plot, grey trace) returns to baseline much more rapidly than that of a slow one (left plot, grey trace). In **a**, a small cooling stimulus and the response it elicited in a representative cell are normalized and overlaid to facilitate comparison of dynamics (note that the sign of cooling stimuli is inverted for clarity, $\Delta T \approx -3^\circ\text{C}$, responses are $\Delta F/F$ recorded at cell bodies, stimuli and responses are recorded synchronously; see Methods for details). The decay constant of each response is shown next to each plot (τ , defined as

the time from peak to $1/e \sim 37\%$ of peak value). **b**, Decay constants of slow and fast cold tPN responses to a range of cold stimuli are systematically different (box-plot: red line, median; blue box, 25th and 75th percentiles; whiskers, data range; $P < 0.001$, t -test; $\Delta T = 2.91 \pm 0.75$, mean \pm s.d.; $n = 16$ stimuli, $n = 7$ slow cells, 9 fast cells). In **c**, a larger cooling stimulus and the response it elicited in a representative slow-adapting and fast-adapting cell are shown side-by-side to facilitate comparison. Note that the slow-adapting cell displays a persistent response that lasts as long as the cooling stimulus. In contrast, as a consequence of faster decay, the response of a 'fast-adapting' cell returns to the pre-stimulus baseline when the temperature is still changing.

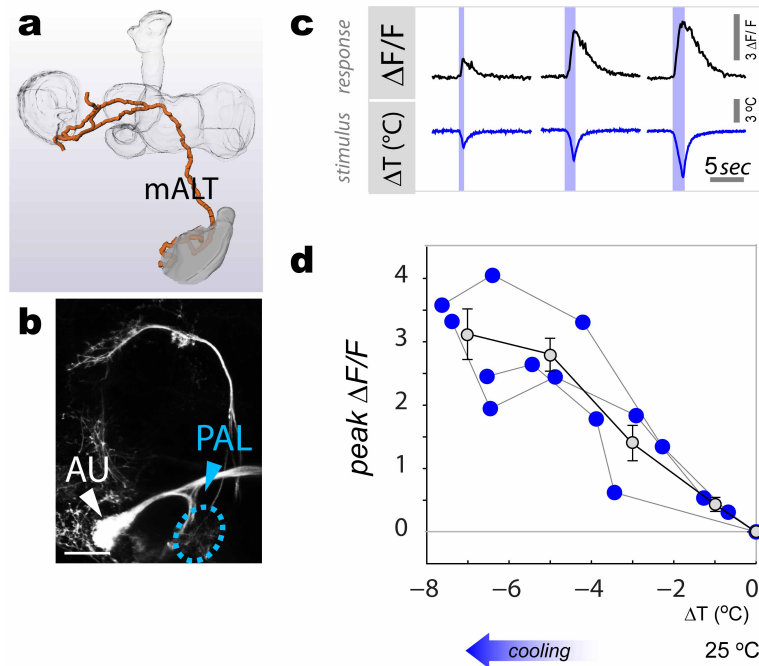


Extended Data Figure 5 | ON and OFF responses amidst a complex temperature stimulus. 'Fast-adapting cold' (purple trace) and 'fast-adapting hot' cells (green trace) track remarkably well a dynamic temperature transient, displaying alternating ON and OFF responses. The cold cell type (purple trace) displays a significantly delayed response whenever a cooling phase is preceded by heating. Notably, cooling responses following a heating step begin when the temperature has nearly returned to baseline (red arrowheads referred to as OFF responses; see main text). Shown here is a representative single stimulus and corresponding response traces simultaneously recorded from adjacent cell bodies; genotype R95C02/UAS-GCaMP6m.



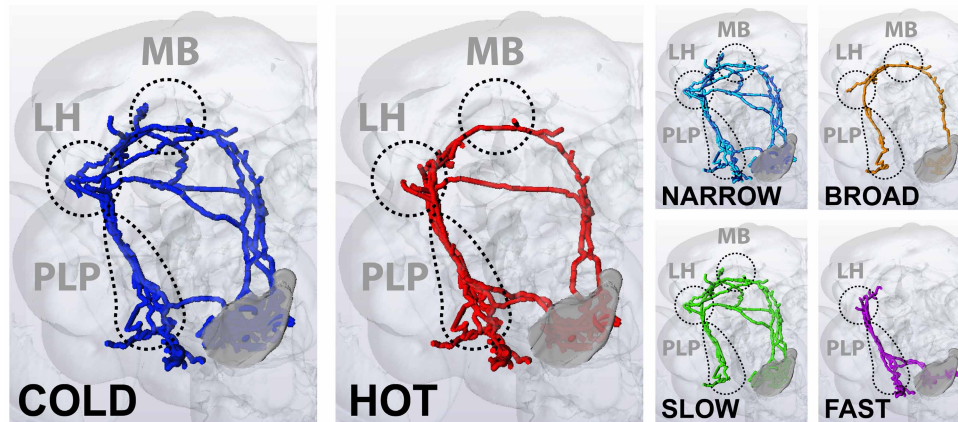
Extended Data Figure 6 | ON and OFF responses of fast tPNs are not limited to a defined range of absolute temperature. To better understand the nature of ON and OFF responses, we tested whether the dynamic activity of fast-adapting cells was limited to a defined range of absolute temperatures by delivering defined heating and cooling stimuli from a range of baseline temperatures. **a**, ON responses from cold (purple traces) and hot (green traces) cells recorded from a 25 $^{\circ}\text{C}$ baseline and from a higher (29 $^{\circ}\text{C}$) or lower (20 $^{\circ}\text{C}$) baseline, respectively (shown are representative single stimuli and corresponding response traces recorded from cell bodies). **b**, **c**, The activity of fast tPNs is not restricted to a specific temperature band as evident by plotting the absolute temperature corresponding to response onset (see arrows in **a**) against the baseline temperature for fast-adapting cold (**b**) and fast-adapting hot (**c**) cells, challenged with a range of hot and cold stimuli (ΔT 2–7 $^{\circ}\text{C}$; cooling $\Delta T = -4.7 \pm 1.9$ $^{\circ}\text{C}$, heating $\Delta T = 4.6 \pm 2$ $^{\circ}\text{C}$, mean \pm s.d.; $n > 10$ cells/ >10 animals per cell type). In **b**, **c**, filled purple and green dots represent ON responses while empty circles represent OFF responses. Owing to their specific delay, the OFF responses of cold cells start at the very end of a hot stimulus (that is, when the temperature has nearly returned to baseline

$T_{\text{onset}} - T_{\text{baseline}} = 0.68 \pm 0.43$ $^{\circ}\text{C}$, mean \pm s.d.); in contrast, hot cells respond at or near the start of the heating phase that follows a cold stimulus (see below) and, as a result, systematically at a lower temperature than baseline ($T_{\text{onset}} - T_{\text{baseline}} = -2.97 \pm 1.15$ $^{\circ}\text{C}$, mean \pm s.d.; note that this value is significantly different than the one calculated for cold cells, $P < 0.001$, t -test). **d–h**, Fast-adapting hot cells innervating the lateral pathway display OFF responses whose delay varies by cell. **d**, **e**, A pair of adjacent hot-activated (green arrow and trace) and cold-activated (purple arrow and trace) tPNs innervating the lateral pathway (3D reconstruction in **d**, cell bodies in **e**; same cells and colour coding as in main Fig. 3). **f**, The cold-activated tPN responded to cooling stimuli with a rapid ON transient. The hot-activated cell in this pair displayed a calcium transient that was not delayed; that is, started immediately with the heating phase that followed the cooling stimulus (light green trace). **g**, **h**, In contrast, a different fast hot cell type in the same cluster (dark green circle and trace) displayed a delayed response (arrowhead), similar to the delayed OFF response reported for cold-activated tPNs at the end of a heating stimulus (see Fig. 3). Scale bars in **e** and **g**, 5 μm .



Extended Data Figure 7 | Cold-specific tPNs innervating the mALT. **a–d**, A distinct group of cold-activated tPNs follows the mALT and innervates the mushroom body and lateral horn (represented by a 3D reconstruction). **b**, The driver VT26020 is active in three neurons innervating this pathway (also showing prominent expression in auditory centres, AU, see Methods for details; PAL, posterior antennal lobe, the approximate position of the cold glomerulus is circled; genotype = VT26020-Gal4/UAS-CD8:GFP). **c**, Representative calcium-imaging traces from a single cell body displaying robust transients in response to cooling stimuli of increasing magnitude (note that these cells do not show activation in response to heating).

d, Stimulus–response plot showing the responses of this cell type to defined cooling stimuli recorded at the cell body. The responses are proportional to cold stimuli over a wide range of intensities ($R^2 = 0.8$, $P < 0.001$; slope = -0.45 , 95% CI = $[-0.57$ to $-0.33]$; regression and slope statistics were calculated on all data points across individual flies). Each blue dot represents the maximal response of a single cell to a defined stimulus; dots are connected when the responses belong to the same cell, $n = 3$ cells/3 animals. Grey dots represent binned averages \pm s.e.m. (bins = $2^{\circ}C$; genotype: VT26020-Gal4/UAS-GCaMP6m). Scale bar in **b**, $25\ \mu m$.



Extended Data Figure 8 | tPN pathways mediating temperature representation in the fly brain. Summary of the array of tPN pathways mediating temperature representation in the fly brain, their properties and target regions. Multiple pathways carry specific information about stimulus quality (hot and cold) to the mushroom body, lateral horn and PLP. Most

tPNs innervating these pathways are narrowly tuned, responding to either hot or cold stimuli, but a prominent group of cells is capable of responding to both heating and cooling temperature stimuli (broadly tuned). Moreover, response dynamics of tPNs are strikingly different, allowing a classification of tPN responses in slow versus fast adapting.

Extended Data Table 1 | Driver lines used in this study and summary of the properties of the tPNs in which they are active

Pathway	Driver	Dendrites (Denmark)	Terminals (sy:t:GFP)	GRASP	Narrowly Tuned	Broadly Tuned	Speed (τ , AV \pm SD; secs)
mALT	R22c06	PAL	MB, LH, PLP	Hot and Cold	<div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div></div>	Slow (8.5 \pm 4.6)
	VT40053	PAL	MB, LH, PLP	Hot and Cold	<div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
		VT26020	PAL	MB, LH	Cold	<div><div></div><div></div><div></div></div>	-
IALT							
<div>Posterior</div> <div>Anterior</div>	R95c02	PAL	LH, PLP	Hot and Cold	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	-	Fast (1.52 \pm 0.5)
	VT19428	PAL	LH, PLP	Cold	<div><div></div><div></div><div></div><div></div><div></div><div></div></div>	-	
	VT46265	PAL	LH, PLP	Hot	<div><div></div></div>	-	Fast (1.5 \pm 0.6)
	VT60737	PAL	LH, PLP	N.D.	<div><div></div></div>	-	
	t3ALT	R84e08	PAL	LH, PLP	Hot and Cold	<div><div></div><div></div></div>	-
t5ALT	R60h12	PAL	MB	Cold	<div><div></div></div>	-	Slow (6.9 \pm 1.7)
	R30b06	N.D.	MB	N.D.	<div><div></div></div>	-	

Driver lines expressed in specific tPNs cell types organized by pathway. Their nomenclature reflects their origin from either the Janelia Farm FlyLight initiative (R lines) or the ViennaTile project (VT lines). The localization of dendritic arborizations in the PAL was defined by transgenic expression of the dendritic marker DenMark, while putative pre-synaptic terminals in target regions were visualized by synaptotagmin-GFP expression (MB, mushroom bodies; LH, lateral horn; PLP, posterior lateral protocerebrum). The GRASP column summarizes the result of experiments performed with hot- or hot-and-cold LexA drivers (see Extended Data Fig. 2 for details). Hot or cold signifies that GRASP reconstitution was limited to the hot or cold glomerulus. 'Hot and cold' denotes broad GFP reconstitution in the PAL. Coloured spots denote the number of tPNs labelled by each driver and their tuning, such that a blue spot represents a single cell responding to cooling while a red spot represents a cell responding to heating. Purple spots are broadly tuned projection neurons. For tuning, cells are defined as hot- or cold-activated if they reliably respond to heating or cooling with average calcium increases of at least 50% $\Delta F/F$ at the onset of a rapid $\sim 2-4^\circ\text{C}$ ΔT stimulus (min $n > 3$ animals/5 cells) and with no significant calcium increases above noise in response to the opposite stimulus (min $n > 3$ animals/5 cells). In all cases, we confirmed that the effect of temperature was significant (see Methods). 'Speed' refers to the decay kinetics of their responses, and is shown as τ (see text and Methods for details). Note that a number of drivers are active in tPNs with different tuning properties but in no cases in both 'fast-adapting' and 'slow-adapting' cells; for this reason a single τ is shown for each driver, corresponding to responses from a variety of stimuli ($n > 10$). *The R84E08 driver supports very low expression of G-CaMP6m. Responses from these cells were recorded at the PAL and appear to be slow ($\tau \sim 4$ s) but a better driver will be needed to describe in detail the response properties and dynamics of these tPNs.

CetZ tubulin-like proteins control archaeal cell shape

Iain G. Duggin^{1,2}, Christopher H. S. Aylett¹, James C. Walsh^{2,3}, Katharine A. Michie¹, Qing Wang¹, Lynne Turnbull², Emma M. Dawson², Elizabeth J. Harry², Cynthia B. Whitchurch², Linda A. Amos¹ & Jan Löwe¹

Tubulin is a major component of the eukaryotic cytoskeleton, controlling cell shape, structure and dynamics, whereas its bacterial homologue FtsZ establishes the cytokinetic ring that constricts during cell division^{1,2}. How such different roles of tubulin and FtsZ evolved is unknown. Studying Archaea may provide clues as these organisms share characteristics with Eukarya and Bacteria³. Here we report the structure and function of proteins from a distinct family related to tubulin and FtsZ, named CetZ, which co-exists with FtsZ in many archaea. CetZ X-ray crystal structures showed the FtsZ/tubulin superfamily fold, and one crystal form contained sheets of protofilaments, suggesting a structural role. However, inactivation of CetZ proteins in *Haloferax volcanii* did not affect cell division. Instead, CetZ1 was required for differentiation of the irregular plate-shaped cells into a rod-shaped cell type that was essential for normal swimming motility. CetZ1 formed dynamic cytoskeletal structures *in vivo*, relating to its capacity to remodel the cell envelope

and direct rod formation. CetZ2 was also implicated in *H. volcanii* cell shape control. Our findings expand the known roles of the FtsZ/tubulin superfamily to include archaeal cell shape dynamics, suggesting that a cytoskeletal role might predate eukaryotic cell evolution, and they support the premise that a major function of the microbial rod shape is to facilitate swimming.

Many archaea produce FtsZ that appears to function in cell division^{4–8}. However, unlike bacteria, archaeal genomes frequently contain additional genes belonging to the FtsZ/tubulin superfamily⁹. These genes are abundant in the haloarchaea, which dominate hyper-saline lakes globally¹⁰ and are generally noted for their unusual flattened cell morphologies. Examples include rectangular prisms^{11,12}, triangles¹³, and pleomorphic cells seen in the model organism *H. volcanii*¹⁴. The molecular basis of cell shape in haloarchaea is unknown.

The largest group of archaeal FtsZ/tubulin proteins clearly distinguishable from the FtsZ and tubulin families is a deeply-rooted family

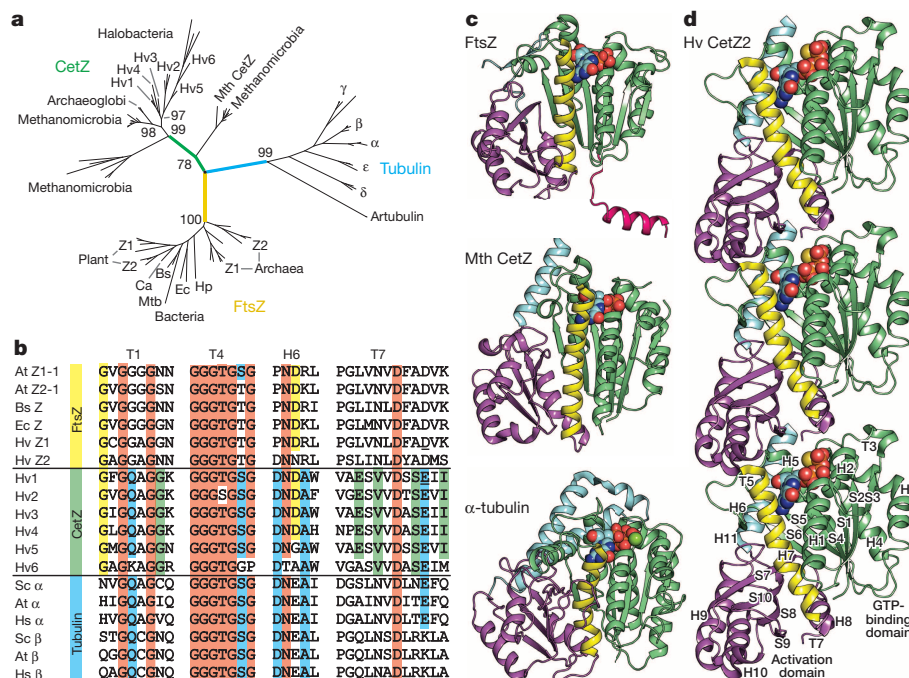


Figure 1 | CetZ proteins comprise a distinct tubulin/FtsZ family.

a, Phylogram of the tubulin superfamily across the domains of cellular life. The major taxa containing identified CetZ family members are shown. Bootstrap support is shown for selected branches. *H. volcanii* CetZ1–6 (Hv1–6) and other CetZ sequences cluster into orthologous groups. Mth CetZ belongs to a non-canonical group that branches from near the centre of the tree. Two archaeal tubulins (Artubulins) identified in Thaumarchaeota (ref. 16) are indicated within the tubulin branch. **b**, Amino acid sequence alignments of some conserved regions. Conservation is indicated by red shading (universal), yellow (primarily shared in FtsZ and CetZ), blue (CetZ and tubulin), or green (CetZ). T7 residues mutated to alanine in FtsZ1 (D250) and CetZ1 (E218) are

underlined. **c**, Structural comparison of FtsZ (PDB 1FSZ), Mth CetZ, and tubulin (PDB 1JFF). **d**, *H. volcanii* CetZ2-GTP-γS crystal structure shows a straight and untwisted protofilament. The C-terminal α-helix (H11, cyan) superimposes well with the equivalent helix in straight tubulin filaments, although differs from the disordered C terminus consistently observed in FtsZ. Secondary structural helices (H), sheets (S) and turns (T) are numbered sequentially. Species abbreviations: Mth, *Methanosaeta thermophila*; Hv, *Haloferax volcanii*; Ca, *Cyanobacterium aponinum*; Ec, *Escherichia coli*; Bs, *Bacillus subtilis*; Mtb, *Mycobacterium tuberculosis*; Hp, *Helicobacter pylori*; At, *Arabidopsis thaliana*; Sc, *Saccharomyces cerevisiae*; Hs, *Homo sapiens*.

¹Medical Research Council Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, UK. ²The iThree Institute, University of Technology Sydney, New South Wales 2007, Australia.

³School of Physics, University of New South Wales, Sydney, New South Wales 2052, Australia.

from the phylum Euryarchaeota that we named CetZ (formerly annotated 'FtsZ³⁹', or 'FtsZ type 2'¹⁵), after the prototypical member described here: 'cell-structure-related euryarchaeota tubulin/FtsZ homologue 1' (CetZ1, Fig. 1a). Multiple sequence alignments showed that CetZ family members contain a mosaic of tubulin-like and FtsZ-like amino acid residues; several of the most conserved residues involved in GTP binding and hydrolysis are identical in CetZ and tubulin (Fig. 1b, blue shading). In some other regions there is stronger sequence similarity between CetZ family members and FtsZ family members, leading to a closer apparent relationship of these two families^{9,16} (Fig. 1a, b).

To further investigate these relationships, we solved the crystal structures of *H. volcanii* CetZ1 (HVO_2204) and CetZ2 (HVO_0745), from separate orthologous groups, and a divergent CetZ from *Methanosaeta thermophila* (Extended Data Table 1). The structures showed the FtsZ/tubulin core fold, with a carboxy-terminal helical extension (H11) like tubulin, but without the amino-terminal helical extension seen in FtsZ proteins (Fig. 1c). CetZ2 crystallized as a protofilament with GTP- γ S that showed subunit interaction geometry, spacing (43 Å subunit repeat) and surface area (941 Å²) consistent with archetypal tubulin and FtsZ protofilaments (Fig. 1d). Furthermore, the protofilaments were arranged as 2D sheets, showing lateral interactions between protofilaments consistent with those of FtsZ and tubulin (Extended Data Fig. 1). These CetZ structures suggest a possible cytoskeletal or cytomotive¹ function.

To search for functions of the CetZ family, we individually deleted the six *cetZ* open reading frames identified in the *H. volcanii* genome (*AcetZ1*; Fig. 2a and Extended Data Table 2). No differences in growth rate or cell size were detected in the six Δ *AcetZ* strains compared to the parent strain (H98), indicating that CetZ proteins are not individually required for cell division (Fig. 2b). To investigate possible functional redundancy amongst the multiple CetZ family members, we designed a dominant-inhibitory point mutation in CetZ1, the most conserved of the *H. volcanii* CetZ family (close orthologues of CetZ1 were found throughout the class Halobacteria; Extended Data Fig. 2). This mutation (CetZ1(E218A), Fig. 1b) was based on well characterized mutants of tubulin and FtsZ that block GTPase-dependent filament disassembly, forming hyper-stable filaments that severely disrupt function^{17–19}. Expression of the equivalent mutant of *H. volcanii* FtsZ1 (D250A, Fig. 1b) resulted in a very heterogeneous cell-size distribution, including many extensively overgrown cells—the hallmark of a major cell division defect (Fig. 2c, d). Conversely, no division defect was detected during CetZ1(E218A) expression in *H. volcanii* *AcetZ1* (Fig. 2c), despite the strong dominant-inhibitory behaviour of this mutant in other functions described below. Therefore, CetZ1 is not required for cell division, whereas FtsZ1 plays an important role.

Motility assays revealed a swimming defect in *H. volcanii* *AcetZ1*; no substantial migration in soft agar medium was observed after 5 days of incubation, whereas normal expanding halos of motile cells were observed in strains carrying single deletions of the other five *cetZ* genes (Fig. 2e). Extended incubation showed that *H. volcanii* *AcetZ1* was motile, with a significantly reduced rate compared to the wild type (Fig. 2f). Both strains also showed exclusion zones between adjacent halos (Fig. 2f and Extended Data Fig. 3a, b), indicating repulsive chemotaxis or sibling colony inhibition²⁰. These results suggest that the *AcetZ1* mutation primarily affects swimming speed.

To confirm a role for CetZ1 in motility, CetZ1(E218A) was expressed under control of the tryptophanase gene promoter on the plasmid pTA962 (ref. 21) in an otherwise wild-type *cetZ* strain (H98). By using increasing tryptophan concentrations, chosen to give dose-responsive expression²², we observed a corresponding reduction in motility compared to the control (Fig. 2g). This dominant-inhibitory effect of CetZ1(E218A) confirms that CetZ1 has an important role in motility.

In Hv-Ca liquid medium, *H. volcanii* exhibited the irregular-plate (or 'disc-shaped'¹⁴) morphology (Fig. 2d). However, we discovered that cells withdrawn from the leading edge of Hv-Ca soft-agar motile halos were rod-shaped (Fig. 2h, see also Supplementary Video 1). The frequency of rod cells decreased markedly towards the centre of the halo, where

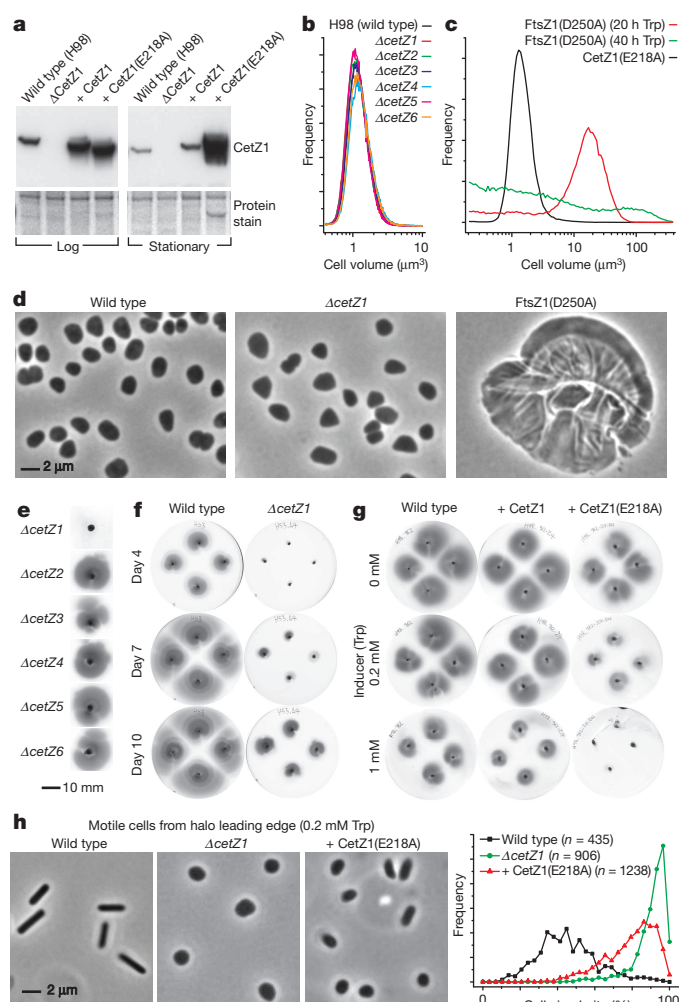


Figure 2 | *H. volcanii* CetZ1 is essential for efficient swimming and rod-cell development, but not cell division. **a**, Western blot detection of CetZ1 in mid-log- and stationary-phase cultures (in the rich medium, Hv-YPC + 1 mM Trp). Ponceau S pre-staining was used as a loading control (bottom). CetZ1(E218A) (in strain ID59 + pTA962-*cetZ1*(E218A)) builds up strongly in non-dividing (stationary phase) cells compared to wild-type CetZ1 (ID59 + pTA962-*cetZ1*), which shows net depletion in the stationary phase, consistent with the notion that CetZ1(E218A) polymer hyper-stability provides increased protection from proteolysis. **b**, Cell volume histograms for the six H98*AcetZ* strains during mid-log growth in Hv-Ca medium. **c**, Cell volume histograms (log-spaced bins) from cell populations producing CetZ1(E218A) (ID59 + pTA962-*cetZ1*(E218A), mid-log growth in Hv-Ca + 1 mM Trp), or FtsZ1(D250A) (H98 + pTA962-*ftsZ1*(D250A), 20 h and 40 h after addition of 1 mM Trp). **d**, Phase-contrast images of wild-type (H98 + pTA962), Δ *AcetZ1* (ID59 + pTA962) and FtsZ1(D250A) (H98 + pTA962-*ftsZ1*(D250A), 40 h Trp) strains. **e**, Motility assays of H98*AcetZ* strains (Hv-Ca + 0.3% agar, day 5). **f**, Time-course motility assays using an alternative genetic background (H53), with four spaced inoculations per plate. The same results were obtained with the H98 background (not shown). **g**, Motility assays (day 5) of wild-type (H98 + pTA962), CetZ1 (H98 + pTA962-*cetZ1*) and CetZ1(E218A) (H98 + pTA962-*cetZ1*(E218A)) strains with the indicated Trp concentrations. High concentrations of Trp partially inhibit growth and motility in all strains. **h**, Cells withdrawn from within 2 mm of the halo edge were viewed by phase-contrast microscopy; cell circularity (right) was measured by automated analysis of the images. Leading-edge wild-type cells had a mean width of $0.720 \pm 0.013 \mu\text{m}$ (95% CI), s.d. = $0.094 \mu\text{m}$, indicating specific regulation of rod-cell width. Frequency in **b**, **c** and **h** is given in arbitrary units.

very few rods were seen (Extended Data Fig. 3d–f). In contrast, samples from the low-motility *AcetZ1* halos were devoid of rods (Fig. 2h). Furthermore, moderate expression of CetZ1(E218A), which reduced motility compared to wild-type CetZ1 (Fig. 2g, 0.2 mM Trp), produced

much more rounded cells at the leading edge of the halo (Fig. 2h). CetZ1 is therefore essential for the development of a rod-shaped cell type required for efficient swimming.

To determine whether CetZ1 directly controls cell-shape changes, wild-type and mutant proteins were overproduced using a plasmid during growth in liquid Hv-Ca. Overproduction of CetZ1 (by two- to threefold) caused a rod morphology in a substantial fraction of cells compared to the control (Fig. 3a, b). During CetZ1(E218A) expression, rods were not detected, and the cells instead formed rather jagged, blocky shapes, exhibiting short stalks and other protrusions (Fig. 3a, c). For both wild-type and CetZ1(E218A) proteins, the extent of the phenotypes was sensitive to the level of production (Extended Data Fig. 4). CetZ1 therefore plays a pivotal role in controlling cell shape.

To examine cell shape and protein subcellular localization, cells producing additional FtsZ1 or CetZ1 tagged with green fluorescent protein (GFP) were visualized. FtsZ1-GFP showed mid-cell bands (rings), which contracted with division furrows, in plates and motile rods (Fig. 4a, b). In plates, CetZ1-GFP was seen throughout the cytoplasm, and localized as spots, short filaments (or patches) of variable intensity at or near the envelope, particularly in regions of relatively high curvature at the cell edges, and at mid-cell and the division furrow (Fig. 4c and Supplementary Video 2). In motile rods, CetZ1-GFP localized at or near one or both poles, as spots, short filaments or end-caps, and at mid-cell in some (Fig. 4d). These patterns were recapitulated with monomeric GFP(A206K) and mCherry tags (Extended Data Fig. 5).

Similar to previous FtsZ and tubulin GFP-fusion proteins^{23,24}, CetZ1-GFP was not fully functional, as its overproduction did not cause rod morphology like untagged CetZ1 (Extended Data Fig. 6a, b). However, overproduced CetZ1-GFP formed dynamic filaments in some cells, typical of the tubulin family (Supplementary Video 3). Furthermore, moderate expression of CetZ1(E218A)-GFP inhibited rod formation (Extended Data Fig. 6d–f), indicating that the fusion protein interacts with native CetZ1.

Cells producing CetZ1(E218A)-GFP showed very low cytoplasmic fluorescence and intense patches of localization at the cell envelope, predominantly at the regions of high curvature caused by the E218A mutation, such as V-grooves (inward), corners and stalks (Fig. 4e, Extended Data Fig. 6a, g, h and Supplementary Video 4). Consistent with

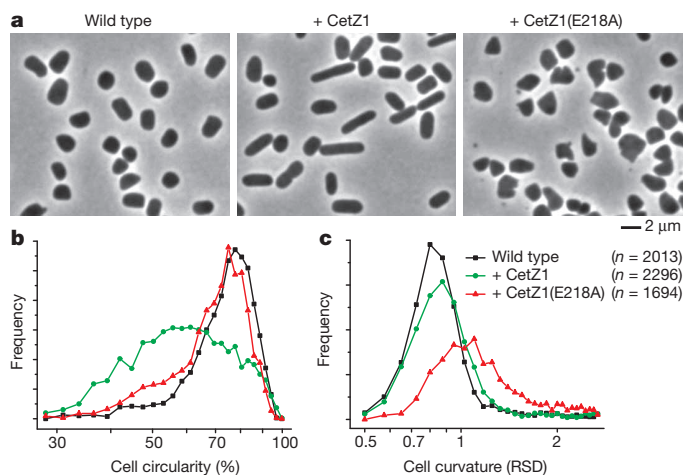


Figure 3 | CetZ1 overproduction changes *H. volcanii* cell shape. Mid-log cultures overproducing the indicated proteins (in Hv-Ca + 2 mM Trp) were sampled for phase-contrast microscopy. **a**, Images of wild-type vector-only (H98 + pTA962), CetZ1 overproduction (H98 + pTA962-*cetZ1*) and CetZ1(E218A) overproduction (H98 + pTA962-*cetZ1*(E218A)) strains. (Note: *H. volcanii* cells lie on their larger flat surfaces.) **b**, **c**, Frequency distributions were obtained for the three strains from automated analysis of cell circularity (**b**), and relative standard deviation (RSD) of curvature (**c**). The symbols for data apply to both **b** and **c**. Frequency in **b** and **c** is given in arbitrary units.

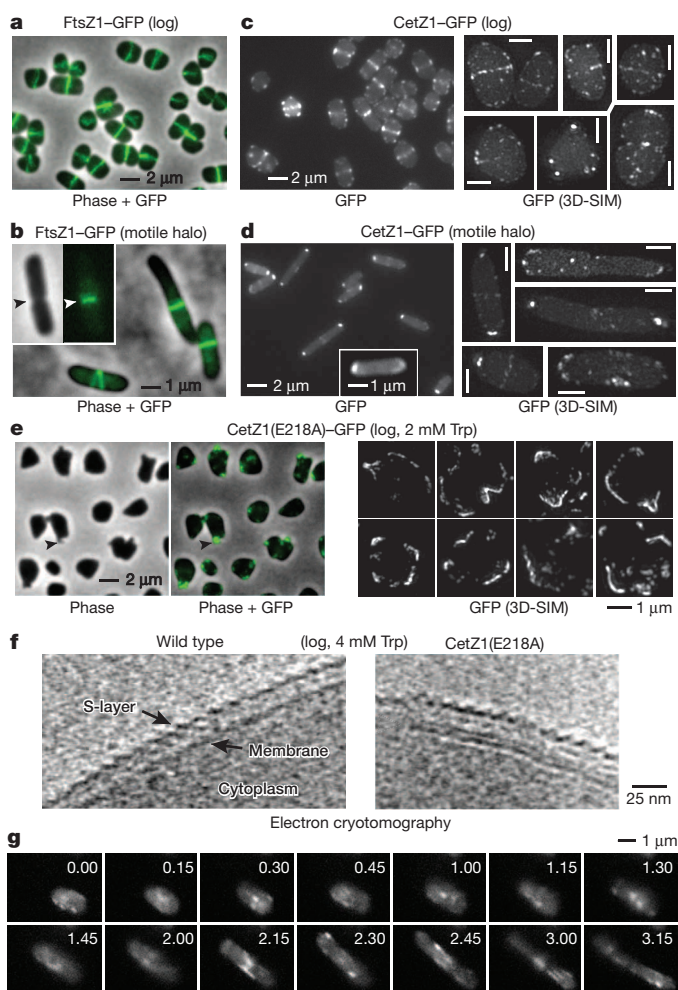


Figure 4 | Subcellular localization of FtsZ1-GFP and CetZ1-GFP in *H. volcanii*. **a**, Fluorescence (green) and phase-contrast (grey) image overlay showing cells producing FtsZ1-GFP (H98 + pIDJL40-*ftsZ1*) sampled during mid-log growth in Hv-Ca + 0.2 mM Trp. **b**, Cells from a motile halo edge (0.2 mM Trp) displaying FtsZ1-GFP, and a cell (inset) showing evidence of division constriction. **c**, CetZ1-GFP in mid-log cells (Hv-Ca + 0.2 mM Trp) visualized by regular fluorescence and three-dimensional structured-illumination microscopy (3D-SIM; xy views of 3D-rendered stacks; all 3D-SIM scale bars, 1 μ m). 3D-SIM z-series showed that the localizations are envelope associated (Supplementary Video 2). **d**, CetZ1-GFP fluorescence in cells from a motile halo edge (0.2 mM Trp), showing detail of an end-capped cell (inset), and 3D-SIM reconstructions (right) showing single or multiple foci, bi-lobed structures, or short filaments. **e**, Phase contrast/fluorescence and 3D-SIM of CetZ1(E218A)-GFP during mid-log growth in Hv-Ca + 2 mM Trp. An example of a cell protrusion and corresponding CetZ1-GFP localization are indicated by the arrowhead. **f**, Ten-nanometre sections from electron cryotomography 3D reconstructions, showing detail of the cell envelope in wild-type (H98 + pTA962) and CetZ1(E218A) cells (H98 + pTA962-*cetZ1*(E218A)), grown in Hv-YPG + 4 mM Trp. Assumed layers of the envelope are indicated, and an extra layer of density on the inner surface of the envelope, which was only clear during CetZ1(E218A) overproduction, is indicated by the arrowhead; 3D segmentation analysis showed that the indicated layer occupies a surface area of 75 nm \times 300 nm. S-layer, surface layer. **g**, Time-lapse imaging of CetZ1-GFP during rod development in Hv-YPG + 0.2 mM Trp (15-min intervals). See also Supplementary Video 8.

this, electron cryotomography sections of cells producing high levels of untagged CetZ1(E218A) showed regions with an additional layer on the inner surface of the envelope (Fig. 4f, Extended Data Fig. 7 and Supplementary Video 5). Time-lapse imaging revealed that CetZ1-GFP localizations were highly dynamic in growing cells, moving in a time frame of seconds to minutes, whereas the bright CetZ1(E218A)-GFP

localizations were hyper-stable (Supplementary Videos 6 and 7). The synthetic effects of CetZ1 (E218A) on cell structure, and its corresponding localization, indicate that it may directly remodel the cell envelope.

Rod development occurred by a transition from a two-dimensional (expansive) growth of plates to a one-dimensional mode in rods, with a concomitant narrowing of cells also clear in some cases; filaments containing CetZ1–GFP were sometimes seen moving along the long axis, particularly at the edges near mid-cell as the rod elongated (Fig. 4g and Supplementary Videos 8, 9 and 10), suggesting that such filaments represent an active form of CetZ1 in rod development.

Altogether, our results show that CetZ1 is a key part of a dynamic cytoskeleton that reshapes the cell during rod development. The structures and functions of *H. volcanii* CetZ1 and CetZ2, which appears to participate in cell-shape control (Extended Data Fig. 8), together with the appearance of homologues in diverse Euryarchaeota (Fig. 1a and Extended Data Fig. 2), suggest that dynamic regulation of cell shape by FtsZ/tubulin-superfamily proteins may pre-date the emergence of eukaryotic cells.

The tubulin-like active-site residues and helix 11 in CetZ family members (Fig. 1) could reflect an early evolutionary divergence of FtsZ and CetZ/tubulin, corresponding to the establishment of a non-cell-division role, followed by a further split of CetZ and tubulin as cell architecture evolved in the Archaea/Eukarya lineage. On the other hand, the slightly higher overall sequence similarities between FtsZ and CetZ would argue that the tubulin-like features of CetZ arose independently of tubulin for similar functional purposes.

The development of the rod-shape to attain normal swimming speed indicates that this is an adaptation to optimize swimming, as generally predicted from theoretical considerations of microbial shapes^{25–27}. Inter-conversion of plate and rod cell types would therefore serve as a stop/go switch that responds to the local environment. We anticipate that *H. volcanii* differentiation and the CetZ family will serve as useful models for cell biology and cytoskeletal evolution.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 29 January; accepted 20 October 2014.

Published online 22 December 2014.

1. Löwe, J. & Amos, L. A. Evolution of cytomotive filaments: the cytoskeleton from prokaryotes to eukaryotes. *Int. J. Biochem. Cell Biol.* **41**, 323–329 (2009).
2. Wickstead, B. & Gull, K. The evolution of the cytoskeleton. *J. Cell Biol.* **194**, 513–525 (2011).
3. Woese, C. R., Kandler, O. & Wheelis, M. L. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl Acad. Sci. USA* **87**, 4576–4579 (1990).
4. Baumann, P. & Jackson, S. P. An archaeobacterial homologue of the essential eubacterial cell division protein FtsZ. *Proc. Natl Acad. Sci. USA* **93**, 6726–6730 (1996).
5. Löwe, J. & Amos, L. A. Crystal structure of the bacterial cell-division protein FtsZ. *Nature* **391**, 203–206 (1998).
6. Margolin, W., Wang, R. & Kumar, M. Isolation of an *ftsZ* homolog from the archaeobacterium *Halobacterium salinarum*: implications for the evolution of FtsZ and tubulin. *J. Bacteriol.* **178**, 1320–1327 (1996).
7. Poplawski, A., Gullbrand, B. & Bernander, R. The *ftsZ* gene of *Haloferax mediterranei*: sequence, conserved gene order, and visualization of the FtsZ ring. *Gene* **242**, 357–367 (2000).
8. Wang, X. & Lutkenhaus, J. FtsZ ring: the eubacterial division apparatus conserved in archaeobacteria. *Mol. Microbiol.* **21**, 313–320 (1996).

9. Vaughan, S., Wickstead, B., Gull, K. & Addinall, S. G. Molecular evolution of FtsZ protein sequences encoded within the genomes of Archaea, Bacteria, and Eukaryota. *J. Mol. Evol.* **58**, 19–29 (2004).
10. Oren, A. Molecular ecology of extremely halophilic Archaea and Bacteria. *FEMS Microbiol. Ecol.* **39**, 1–7 (2002).
11. Burns, D. G. et al. *Haloquadratum walsbyi* gen. nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int. J. Syst. Evol. Microbiol.* **57**, 387–392 (2007).
12. Walsby, A. E. Square bacterium. *Nature* **283**, 69–71 (1980).
13. Takashina, T., Hamamoto, T., Ootai, K., Grant, W. D. & Horikoshi, K. *Haloarcula japonica* sp. nov., a new triangular halophilic archaeobacterium. *Syst. Appl. Microbiol.* **13**, 177–181 (1990).
14. Mullakhanbhai, M. F. & Larsen, H. *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch. Microbiol.* **104**, 207–214 (1975).
15. Marchler-Bauer, A. et al. CDD: conserved domains and protein three-dimensional structure. *Nucleic Acids Res.* **41**, D348–D352 (2013).
16. Yutin, N. & Koonin, E. V. Archaeal origin of tubulin. *Biol. Direct* **7**, 10 (2012).
17. Löwe, J., Li, H., Downing, K. H. & Nogales, E. Refined structure of α -tubulin at 3.5 Å resolution. *J. Mol. Biol.* **313**, 1045–1057 (2001).
18. Richards, K. L. et al. Structure-function relationships in yeast tubulins. *Mol. Biol. Cell* **11**, 1887–1903 (2000).
19. Scheffers, D. J., de Wit, J. G., den Blaauwen, T. & Driessen, A. J. GTP hydrolysis of cell division protein FtsZ: evidence that the active site is formed by the association of monomers. *Biochemistry* **41**, 521–529 (2002).
20. Be'er, A., Florin, E. L., Fisher, C. R., Swinney, H. L. & Payne, S. M. Surviving bacterial sibling rivalry: inducible and reversible phenotypic switching in *Paenibacillus dendritiformis*. *MBio* **2**, e00069–11 (2011).
21. Allers, T., Barak, S., Liddell, S., Wardell, K. & Mevarech, M. Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii*. *Appl. Environ. Microbiol.* **76**, 1759–1769 (2010).
22. Large, A. et al. Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential *cct1* gene. *Mol. Microbiol.* **66**, 1092–1106 (2007).
23. Kimble, M., Kuzmiak, C., McGovern, K. N. & de Hostos, E. L. Microtubule organization and the effects of GFP-tubulin expression in *Dictyostelium discoideum*. *Cell Motil. Cytoskeleton* **47**, 48–62 (2000).
24. Sun, Q. & Margolin, W. FtsZ dynamics during the division cycle of live *Escherichia coli* cells. *J. Bacteriol.* **180**, 2050–2056 (1998).
25. Cooper, S. & Denny, M. W. A conjecture on the relationship of bacterial shape to motility in rod-shaped bacteria. *FEMS Microbiol. Lett.* **148**, 227–231 (1997).
26. Dusenbery, D. B. Fitness landscapes for effects of shape on chemotaxis and other behaviors of bacteria. *J. Bacteriol.* **180**, 5978–5983 (1998).
27. Young, K. D. The selective value of bacterial shape. *Microbiol. Mol. Biol. Rev.* **70**, 660–703 (2006).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank T. Allers for discussion, strains and plasmids, F. Pfeiffer, M. Dyall-Smith, R. Cavicchioli, I. Charles, C. Angstrom and P. Curmi for discussion, J. Maupin-Furlow and D. Sherratt for plasmids, M. Johnson and the UTS Microbial Imaging Facility for technical support, F. Gorrec and S. Köhlmann for help at the MRC-LMB crystallization facility, and the European Synchrotron Radiation Facility and Diamond Light Source for service and support. This work was supported by the Medical Research Council, UK (U105184326 to J.L.) and the University of Technology Sydney, Australia. C.B.W. was supported by the NHMRC, Australia (SRF 571905).

Author Contributions I.G.D., K.A.M. and J.L. initiated the project. I.G.D., C.H.S.A., K.A.M., L.A.A. and J.L. designed the experiments. I.G.D. and E.M.D. carried out molecular phylogeny. C.H.S.A., J.L., L.A.A. and I.G.D. solved or analysed crystal structures. I.G.D. carried out genetic modification and phenotypic analysis. I.G.D., J.C.W., L.T., C.B.W. and E.J.H. carried out light microscopy. J.C.W. and I.G.D. carried out image analysis. Q.W. carried out electron cryotomography. I.G.D. wrote the manuscript, and all authors reviewed the manuscript.

Author Information Crystal structures have been deposited in the Protein Data Bank (PDB) under the accession codes 4B46 (*H. volcanii* CetZ1), 4B45 (*H. volcanii* CetZ2) and 3ZID (*M. thermophila* CetZ). Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to I.G.D. (Iain.Duggin@uts.edu.au).

METHODS

Molecular phylogeny. Protein sequences were retrieved from the National Centre of Biotechnology Information (NCBI) website. PSI-BLAST²⁸ was used to identify *CetZ* sequences using *CetZ1* (HVO_2204) from *H. volcanii* as the query sequence. Protein sequences were aligned using MUSCLE²⁹, and then a phylogenetic tree was generated using Fasttree³⁰.

Protein purification. The open reading frames for *H. volcanii* *CetZ1* (HVO_2204) and *CetZ2* (HVO_0745) were first amplified from genomic DNA and then cloned into the expression plasmid pHis17 (ref. 31). pHis17-Hv*CetZ1* encoded the complete sequence of *H. volcanii* *CetZ1* without additional residues, whereas pHis17-Hv*CetZ2* encoded the complete sequence of *H. volcanii* *CetZ2*, adding six histidine residues at the C terminus. The gene encoding *M. thermophila* *CetZ* (Mthe_0239, Uniprot A0B5R2) was synthesized codon-optimized (Genscript), and cloned into pET28a (pET28a-Mt*CetZ*), including an additional six histidine residues at the C terminus.

Escherichia coli C41 strains harbouring the expression vectors were grown in 12 litres of 2 × YT broth supplemented with 100 µg ml⁻¹ ampicillin or 50 µg ml⁻¹ kanamycin and grown at 37 °C. Isopropyl-β-D-1-thiogalactopyranoside was added to a final concentration of 1 mM (at D₆₀₀ = 1.0), and incubation continued for 18 h at 20 °C. Collected cells were resuspended in 20 ml per litre of bacterial culture of 100 mM Na-HEPES pH 7.0, and then lysed using a cell disruption system (Constant Systems) running at 25 kPSI at 4 °C. The mixture was centrifuged at 45,000g for 30 min at 4 °C. *H. volcanii* *CetZ2* and *M. thermophila* *CetZ* were purified from the supernatant by immobilized metal affinity chromatography (IMAC), using a 5 ml HisTrap HP column (GE Healthcare) running in 100 mM Tris-Cl (pH 8.5), 500 mM KCl, 0–1 M imidazole gradient), followed by ion exchange with a HiTrap Q column HP (GE Healthcare) running in 25 mM Tris-Cl pH 8.5, with a 0–500 mM KCl gradient. Size-exclusion chromatography was then carried out using a HiLoad Sephacryl S200 16/60 column (GE Healthcare), running in 25 mM Tris-Cl pH 8.5, 200 mM KCl, 1 mM EDTA, 1 mM Na₂S₂O₃. High-yielding fractions were pooled and concentrated to 25 mg ml⁻¹ before and after size exclusion chromatography in 20 ml capacity, 30 kDa cut-off centrifugal concentrators (Vivaspin). Untagged *H. volcanii* *CetZ1* was prepared with the same method, except without the IMAC step.

Crystallization. Initial conditions were identified using the high-throughput MRC-LMB crystallization facility³². All *CetZ* crystals were produced in 1 µl protein drops containing 50% protein and 50% (v/v) precipitant solutions. *H. volcanii* *CetZ1* was crystallized in 20% (w/v) polyethylene glycol Mr 6000, 1 M LiCl, 100 mM sodium citrate (pH 4.0). *H. volcanii* *CetZ2* was crystallized in 3–3.5 M ammonium sulphate, 100 mM sodium citrate (pH 5.0). *M. thermophila* *CetZ* crystallized in 30% (v/v) 1,2-propanediol, 20% (v/v) polyethylene glycol (Mr 400), 0.1 M Na-HEPES (pH 7.5). Artificial mother liquor supplemented to 25% (v/v) glycerol was used as a cryo-protectant for all proteins during X-ray diffraction image acquisition.

Crystallographic data collection and structure determination. Diffraction images were collected from single frozen crystals at European Synchrotron Research Facility beamline ID23eh1 (*H. volcanii* *CetZ1* and *M. thermophila* *CetZ*) or Diamond beamline I03 (*H. volcanii* *CetZ2*) at 100 K. The wavelengths used for data collection were 0.9 Å and 0.9795 Å at the respective beamlines. Diffraction images were processed with XDS³³, POINTLESS³⁴ and SCALA³⁵ software. Initial phases were determined by molecular replacement using PHASER³⁶. Models were built manually using MAIN³⁷ and refined using REFMAC³⁸ and PHENIX.REFINE³⁹. Their respective Ramachandran plots were used to validate the structures. The atomic model for *H. volcanii* *CetZ1* placed 99.7% of residues in the most favoured (98%) region; *H. volcanii* *CetZ2* had 98.2%, and *M. thermophila* *CetZ* had 99.1%. No residues from any of the three structures had angles lying in the disallowed (0.2%) regions. The density connecting the helical C-terminal residues of *CetZ2* to the main body of the protein was weak and could not be threaded; therefore, the ordered position of a disconnected C terminus from the adjacent subunit was deposited as a separate chain (chain B). Crystallographic data are given in Extended Data Table 1.

***H. volcanii* media and culture conditions.** Most cultures were grown with Hv-Ca medium⁴⁰ modified to contain an expanded trace-elements solution that improved growth and cell-shape uniformity. The trace-elements stock solution contained, per litre: 5 g EDTA, 0.8 g FeCl₃, 0.05 g ZnCl₂, 0.01 g CuCl₂, 0.01 g CoCl₂, 0.01 g H₃BO₃, 1.6 g MnCl₂, 0.01 g Ni₂SO₄ and 0.01 g H₂MoO₄. The solution was adjusted to pH 7.0 with NaOH, filter sterilized and stored at room temperature. It was diluted 100-fold in media to obtain the working concentration. Where necessary, media was supplemented with auxotrophic requirements⁴⁰, or the indicated concentration of L-tryptophan (Sigma). Hv-YPG medium⁴⁰ was used in some experiments, where indicated, and in genetic modification procedures. For most experiments, cultures were incubated at 45 °C with rotary shaking (150 rpm), and were maintained in log growth (D₆₀₀ < 0.8) for at least 2 days before sampling (at D₆₀₀ = 0.2–0.8), unless otherwise indicated. In addition to preliminary experiments carried out for this

study, representative data displayed are representative of at least two biological replicate experiments.

Strain and plasmid construction. Strain and plasmid construction information and relevant oligonucleotide sequences are given in Extended Data Table 2. Plasmids for controlled expression in *H. volcanii*, containing the *tna* promoter (*Ptna*)²², were based on pTA962²¹. The open reading frame (ORF) encoding the smRS-GFP variant was amplified from pJAM1020 (ref. 41), using overlap extension PCR to remove the internal NdeI site, and cloned between BamHI and NotI of pTA962, creating the vector pIDJL40 (see Extended Data Table 2). The *ftsZ1*(D250A) and various *cetZ* ORFs were PCR amplified from *H. volcanii* DS70 (refs 42, 43) DNA, and then cloned between NdeI and BamHI in both pTA962 and pIDJL40. An mCherry ORF was cloned between BamHI and NotI in pIDJL40-*cetZ1*, replacing *gfp*, to create pIDJL117 (Extended Data Table 2). Point mutations were generated by overlap-extension PCR and fragments were cloned as above, incorporating the following codon substitutions: *ftsZ1*(D250A) (GAC > GCC), *cetZ1*(E218A) (GAG > GCG) and *gfp*(A206K) (GCC > AAG). All PCR-generated fragments were sequence-verified after cloning.

Plasmids for making deletions in *H. volcanii* were constructed by cloning upstream and downstream DNA fragments that flanked each deletion into pTA131 (ref. 40). Two stages were used to generate each *cetZ*-deletion plasmid. In the first stage, the upstream DNA fragment amplified from DS70 (primer pairs US-f/US-r, Extended Data Table 2) was spliced by overlap-extension PCR to the corresponding *Ptna*-*cetZ* fragment that had been amplified with primer PtnaUS-f (CTGGCGAAAGG GGGATGTGCTGCT), the *cetZ* ORF reverse primer, and the expression-plasmid template (Extended Data Table 2). The spliced product, in which the native *cetZ* promoter was swapped for *Ptna*, was cloned between the HindIII–BamHI sites of pTA131. These plasmids were demethylated by passage through *E. coli* C2925 (New England Biolabs), and then used in the ‘pop-in pop-out’ method⁴⁰, to replace the chromosomal *cetZ* promoters with *Ptna*. Transformants were selected as previously described⁴⁰, and then chromosome structures were verified by allele-specific PCR. Preliminary experiments using the resulting strains showed no significant growth defects upon depletion of *cetZ* gene expression by withdrawal of Trp during mid-log growth. In the second stage, the BglII–BamHI (*Ptna*-*cetZ*) fragment in each plasmid was replaced with a downstream fragment (amplified with primers DS-f/DS-r, Extended Data Table 2), to create plasmids for deletion of each *cetZ* ORF, as described above. The resulting strains are listed in Extended Data Table 2.

Western blotting. *CetZ1* rabbit antisera were generated with a synthetic peptide antigen derived from the sequence of the C-terminal region of *CetZ1*: [C]-QDNIEIRQESDSNLETILN-amide (Cambridge Research Biochemicals, UK). Antisera were assayed and verified for specificity (Fig. 2a). *H. volcanii* cells were resuspended in SDS-PAGE sample buffer and then heated (95 °C, 5 min) before analysis by SDS-PAGE and western blotting (10⁻³ primary antisera dilution) using a Criterion electrophoresis system (Bio-Rad) with nitrocellulose membrane (Protran, Whatman) and Chemiluminescence detection (Thermo Scientific) according to standard protocols.

Coulter counting. Culture samples were diluted (10⁻² or 10⁻³) with 18% buffered salt water (BSW) (which included calcium)⁴⁰, and then ~10⁵ cells were analysed with a Multisizer 4 Coulter Counter (Beckman Coulter) in the volumetric mode (100 µl), with a 20 µm aperture tube and 18% BSW as the electrolyte. A 2 µm bead standard (Beckman Coulter), diluted in 18% BSW, was used for prior calibration.

Motility assays. One microlitre of a mid-log culture in Hv-Ca was inoculated into the subsurface of an Hv-Ca + 0.3% (w/v) agar plate (20 ml of media per plate), with up to four inoculations per plate from the same liquid culture. Equal cell densities of the inoculum were used for comparison of multiple strains. Plates were incubated with minimal disturbance in a closed plastic bag at 45 °C for 5 days⁴⁴, unless otherwise indicated, and then images of the plates were acquired using a densitometer/scanner.

Light microscopy. For most microscopy, a 1 µl sample of culture was placed directly onto a slide prepared with a ~170 µm thick, 1% agarose pad containing 18% BSW. A number 1.5 coverslip was then placed on top. Phase-contrast and GFP epifluorescence images were acquired with a Zeiss AxioPlan 2 system with a 1.4 NA phase-contrast objective (Carl Zeiss, Germany).

Time-lapse imaging was performed at 37 °C, with a Nikon Ti with the Perfect Focus system and EMCCD camera (Cascade 1K, Photometrics), or the OMX superresolution system with wide-field deconvolution, described further below. Samples were prepared by placing a 0.5–1 µl sample of growing culture at the centre of a ~0.5-mm-thick gel pad (containing media semi-solidified with 0.2–0.3% agarose or agar) prepared on a coverslip (circular, 8 mm diameter, number 1.5); this was then inverted onto the base of a 35 mm number 1.5 FluoroDish (World Precision Instruments, Inc.), and then 3 ml of liquid media was added to cover the pad assembly. The lid was replaced and the dish was then incubated on the microscope stage (at 37 °C) for 30 min before imaging.

Super-resolution imaging was performed with a V3 DeltaVision OMX 3D-SIM Blaze system (Applied Precision Inc., a GE Healthcare Company, Issaquah, USA).

A solid-state 488 nm wavelength laser provided wide-field illumination and images were captured using a 60× 1.4 NA UPlanSapo oil objective (Olympus, Japan), a standard 488 nm excitation and 500–550 nm emission filter set and a scientific CMOS 512 × 512 pixel 15-bit camera (pco.edge, PCO AG, Kelheim, Germany). For wide-field deconvolution imaging, samples were sectioned using a 250 nm step-size and deconvolved using SoftWorX (Applied Precision, GE). For three-dimensional structured-illumination microscopy (3D-SIM), interference patterns were generated by interfering light beams⁴⁵ and samples were sectioned using a 125 nm z-step size. Raw 3-phase images were then reconstructed as previously described^{46,47}. Deconvolved or reconstructed images-stacks were 3D rendered and presented using IMARIS software (v7.6.4, Bitplane Scientific). Video was prepared using ImageJ⁴⁸.

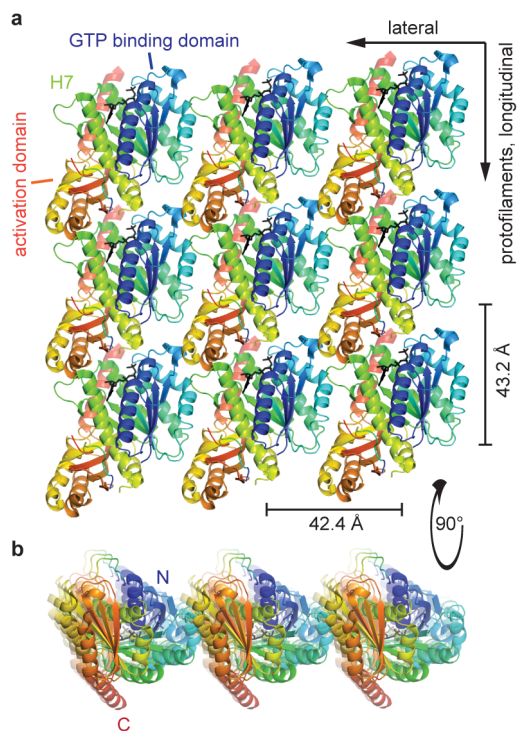
Automated image analysis. Phase-contrast and GFP epifluorescence images were analysed using Mathematica software (Wolfram Research). Phase-contrast images were first binarized by using a bottom-hat transform with a 30-pixel disk matrix for the kernel. Individual objects (cells) were then identified by using the 'Morphological Components' function. Background objects were excluded by three sanity checks using the 'Component Measurements' function: (1) the minimum cell size was 200 pixels (1 pixel = 60 nm) with a bounding disk radius of 4 pixels; (2) fragmented objects were excluded if the ratio of areas calculated by the pixel count and the 'Area Measurement' parameter lay outside 0.95–1.05; (3) cells had to contain no holes. The circularity of each cell was calculated as the bounding disk coverage, that is, the percentage area occupied by a cell within the minimal circle that completely contains the cell outline, by using the 'Component Measurements' function. To measure cell curvature, a curve was parametrically fitted to the outline of each cell by two Fourier expansions up to the 10th order mode; the relative standard deviation (RSD) of curvature was calculated from this line.

To measure fluorescence association with the cell envelope, epifluorescence images were overlaid with cell outlines from each corresponding phase-contrast image. Fluorescence was then integrated along the normal of the parametric fit of the outline for 3 pixels towards the centre of the cell (Extended Data Fig. 7). In cells displaying localization of fluorescence (RSD > 0.5), the mean fluorescence in areas of high curvature (>2 standard deviations above the mean) was compared to the mean fluorescence intensity of the full perimeter and to the mean of the whole cell. The mean percentage difference between each set of two values across all cells was calculated.

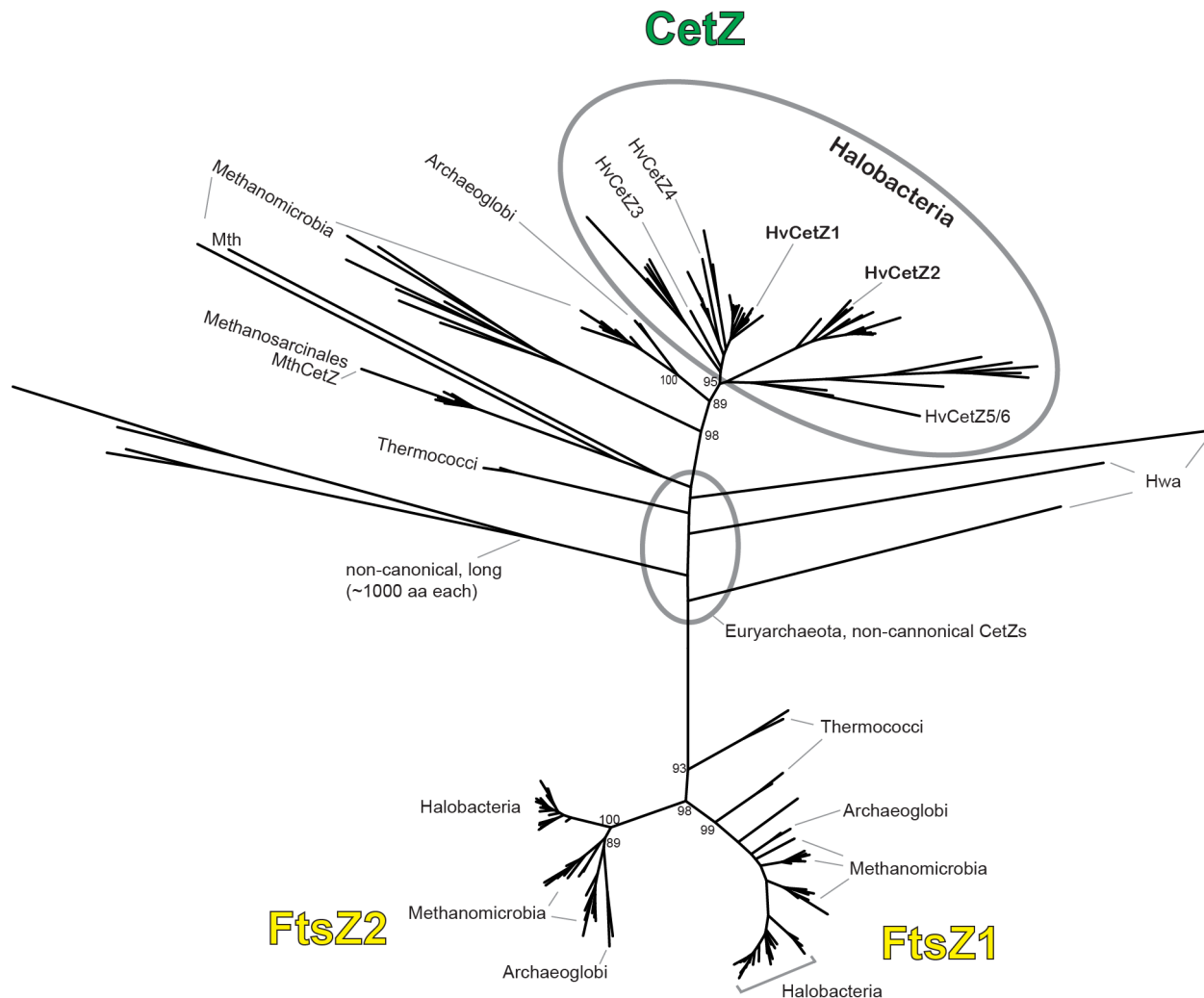
Electron cryotomography. *H. volcanii* were grown in Hv-YPC medium + 4 mM Trp. Mid-log cells were concentrated 10–20-fold by centrifugation and resuspension, mixed with 10-nm protein-A-coated colloidal gold particles (Sigma), and then plunge-frozen on Quantifoil R3.5/1 holey carbon-coated transmission electron microscopy grids using a Vitrobot (FEI Company, USA). The grids were then stored under liquid nitrogen. Imaging was performed at 300 kV on an FEI Tecnai G2 Polara transmission electron microscope equipped with a Gatan energy filter (zero-loss peak mode and 20 eV slit width) and a post-GIF 4 × 4 k Gatan Ultrascan camera binned to 2 × 2 k (Gatan, USA). Tilt series were collected around a single axis at 1.5° increments between ± 42° at a pixel size of 0.58 nm or 0.7 nm (at the specimen level) using SerialEM software⁴⁹. The defocus was set at −12 µm and the cumulative dose was <80 e Å^{−2}. Tomographic reconstructions were calculated using the IMOD tomography package⁵⁰.

28. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).

29. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).
30. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
31. Miroux, B. & Walker, J. E. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J. Mol. Biol.* **260**, 289–298 (1996).
32. Stock, D., Perisic, O. & Löwe, J. Robotic nanolitre protein crystallisation at the MRC Laboratory of Molecular Biology. *Prog. Biophys. Mol. Biol.* **88**, 311–327 (2005).
33. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
34. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D* **62**, 72–82 (2006).
35. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
36. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
37. Turk, D. MAIN software for density averaging, model building, structure refinement and validation. *Acta Crystallogr. D* **69**, 1342–1357 (2013).
38. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
39. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
40. Allers, T., Ngo, H. P., Mevarech, M. & Lloyd, R. G. Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl. Environ. Microbiol.* **70**, 943–953 (2004).
41. Reuter, C. J. & Maupin-Furlow, J. A. Analysis of proteasome-dependent proteolysis in *Haloferax volcanii* cells, using short-lived green fluorescent proteins. *Appl. Environ. Microbiol.* **70**, 7530–7538 (2004).
42. Hartman, A. L. *et al.* The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS ONE* **5**, e9605 (2010).
43. Wendoloski, D., Ferrer, C. & Dyall-Smith, M. L. A new simvastatin (mevinolin)-resistance marker from *Haloarcula hispanica* and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology* **147**, 959–964 (2001).
44. Tripepi, M., Imam, S. & Pohlschroder, M. *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* **192**, 3093–3102 (2010).
45. Strauss, M. P. *et al.* 3D-SIM super resolution microscopy reveals a bead-like arrangement of FtsZ and the division machinery: implications for triggering cytokinesis. *PLoS Biol.* **10**, e1001389 (2012).
46. Gustafsson, M. G. L. *et al.* Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. *Biophys. J.* **94**, 4957–4970 (2008).
47. Schermelleh, L. *et al.* Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. *Science* **320**, 1332–1336 (2008).
48. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671–675 (2012).
49. Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
50. Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J. Struct. Biol.* **116**, 71–76 (1996).
51. Landgraf, D., Okumus, B., Chien, P., Baker, T. A. & Paulsson, J. Segregation of molecules at cell division reveals native protein localization. *Nature Methods* **9**, 480–482 (2012).
52. Zacharias, D. A., Violin, J. D., Newton, A. C. & Tsien, R. Y. Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* **296**, 913–916 (2002).
53. Shaner, N. C. *et al.* Improved monomeric red, orange and yellow fluorescent proteins derived from *Discosoma* sp. red fluorescent protein. *Nature Biotechnol.* **22**, 1567–1572 (2004).

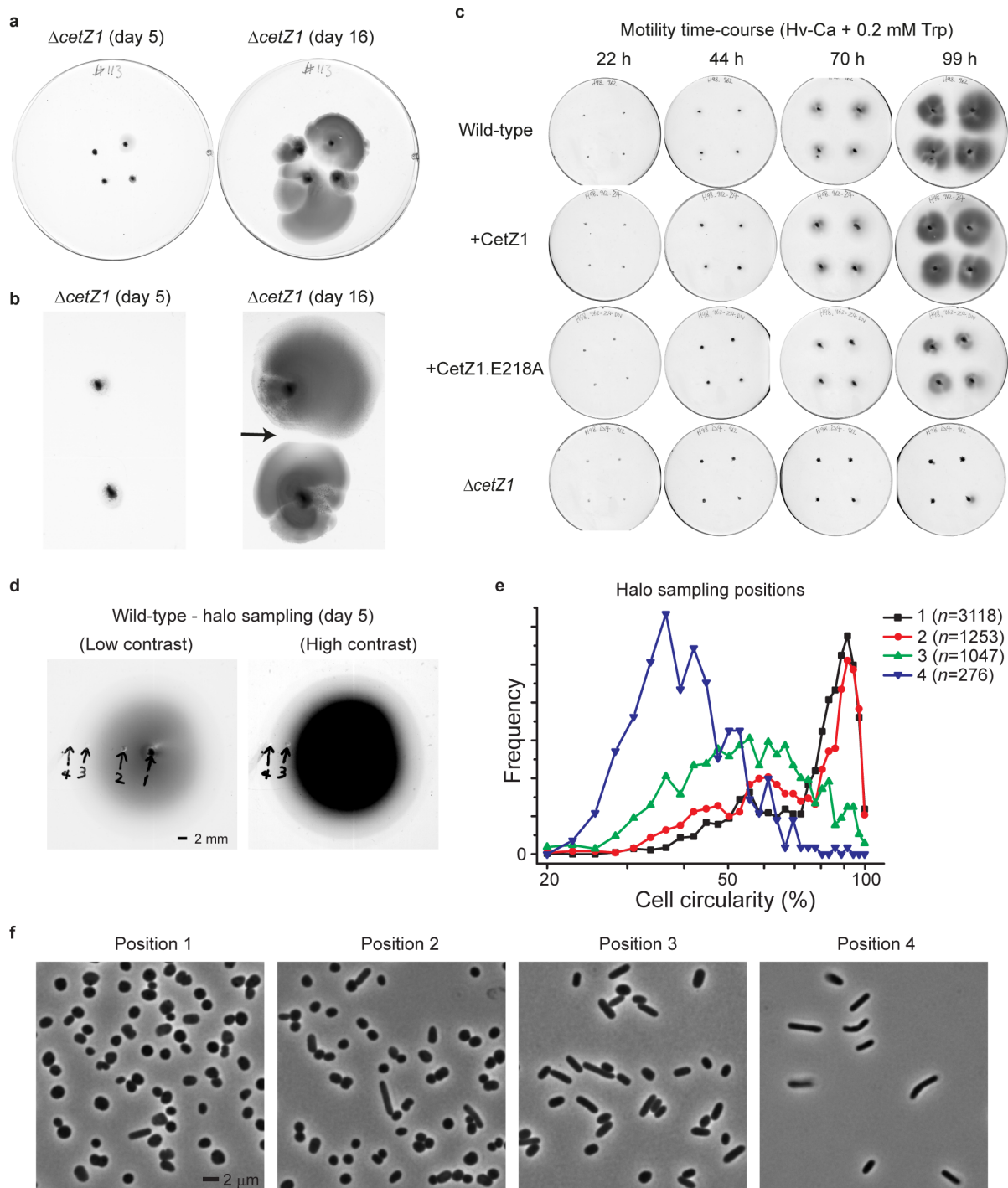


Extended Data Figure 1 | CetZ2-GTP- γ S crystals contain protofilaments and 2D sheets. **a**, Expansion of crystal symmetry (PDB 4B45) shows lateral association of tubulin/FtsZ-like protofilaments. The lateral repeat distance of 42.4 Å is slightly shorter than the longitudinal distance (43.2 Å). There is similar packing to the lattice of tubulin dimers in the microtubule wall, although a specific M-loop facilitates the lateral contact in tubulin. CetZ2 shows less intimate contacts between GTPase domains (green) in the protofilament compared to tubulin. **b**, 90°-rotated view along the protofilaments. Protofilaments are arranged in the same orientation, creating two different surfaces of the sheet. The H11 helix and C termini are located on the lower surface as shown, and would have the potential to interact with other molecules or surfaces (as seen with the C-terminal tail of FtsZ).



Extended Data Figure 2 | Tubulin-superfamily proteins in Euryarchaeota that contain CetZ. CetZ1 from *H. volcanii* (HVO_2204, Uniprot D4GVD7) was used in PSI-BLAST searches to identify tubulin-superfamily members in Euryarchaeota. A representative genome from all genera determined to contain a CetZ or non-canonical CetZ was selected (if available), and then all identified superfamily members from these genomes were aligned for generating the phylogram. Bootstrap support is shown for selected branches. The positions of the *H. volcanii* CetZ proteins are indicated (HvCetZ). All species identified to contain a CetZ contained at least one FtsZ; most contained

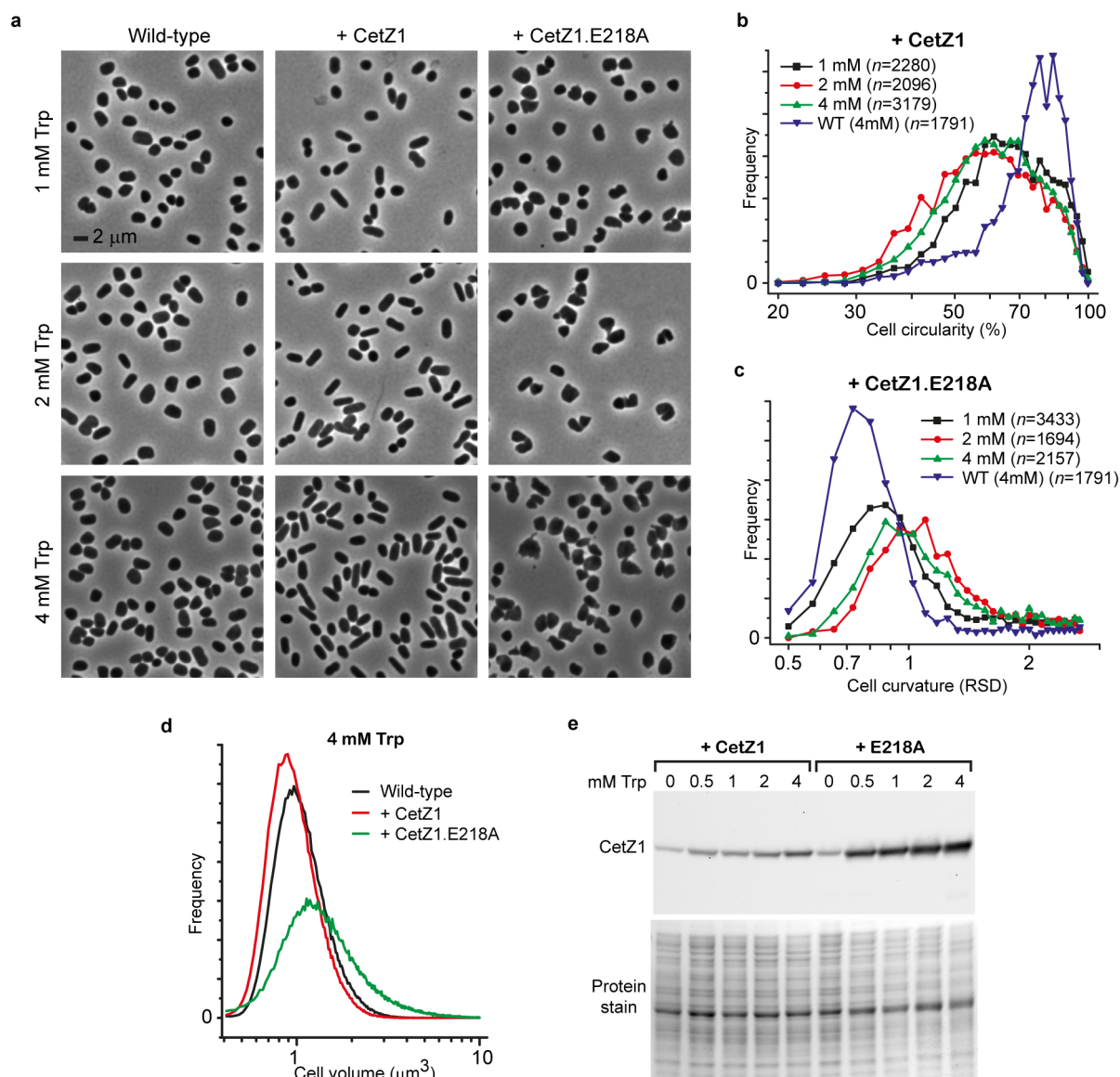
two FtsZ proteins from distinct sub-families. CetZ1 orthologues were identified in all species of Halobacteria analysed, and typically shared ~80% sequence identity. CetZ2 orthologues were identified in many of these. A number of highly diverse sequences containing recognizable tubulin-superfamily signature motifs (Fig. 1b) were identified in Euryarchaeota, including the classes Thermococci, Methanomicrobia and Halobacteria (notably, in the rectangular prism-shaped *Haloquadratum walsbyi* (Hwa)). These were difficult to classify robustly; however, most were found to branch from near the base of the CetZ-family bough and were referred to as non-canonical CetZs.



Extended Data Figure 3 | The role of CetZ1 in *H. volcanii* motility.

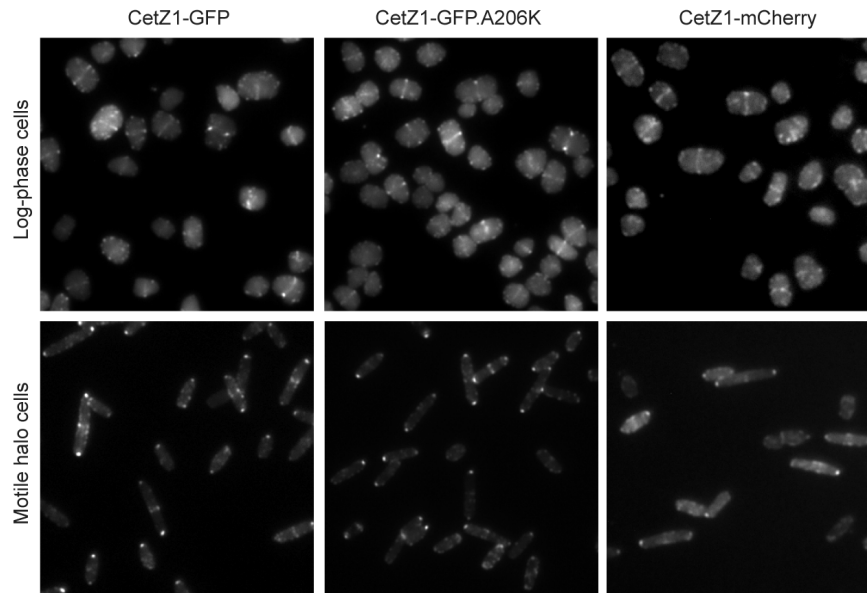
a, b, Motility assays of the $\Delta cetZ1$ strain (ID59 + pTA962) were performed with closely-spaced (**a**) or standard-spaced (**b**) inoculation sites. Exclusion zones were observed between halos after extended incubation (16 days). Irregularities in halo structures were seen in wild-type and mutant halos, which may be related to colony exclusion (repulsion or inhibition) between local clouds of motile cells. **c**, Motility agar plates (containing 0.2 mM Trp), were imaged once daily over 4 days after inoculation. Strains containing wild-type CetZ1 (H98 + pTA962 and H98 + pTA962-*cetZ1*) show an abrupt development of

motile halos from day 3. Strains containing *cetZ1* mutations (H98 + pTA962-*cetZ1*(E218A) and ID59 + pTA962 ($\Delta cetZ1$)), were strongly inhibited in motility. **d**, A halo of wild-type *H. volcanii* (H98 + pTA962) is shown in low and high contrast, showing the high-motility fainter outer zone and the denser population of cells growing more centrally. Samples were withdrawn from the indicated positions (1–4) of the motility halo for microscopy. **e, f**, Cell circularity distributions (**e**) were then generated from automated computer analysis of images of cells (**f**) sampled at the four positions.



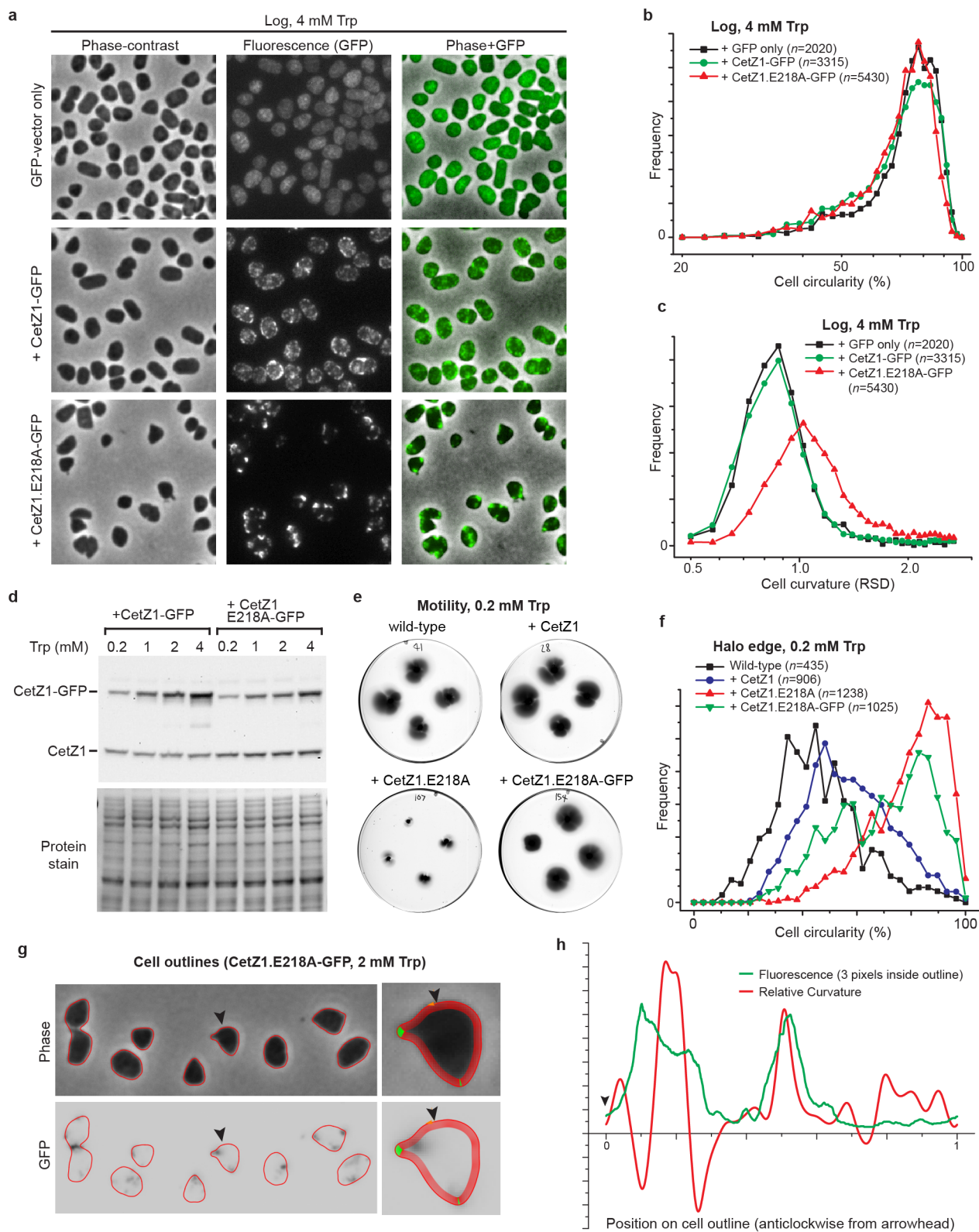
Extended Data Figure 4 | *H. volcanii* cell shape is sensitive to CetZ1 and CetZ1(E218A) concentration. **a**, Phase-contrast images of wild-type (H98+pTA962), CetZ1-overproduction (H98+pTA962-*cetZ1*) and CetZ1(E218A)-overproduction (H98+pTA962-*cetZ1*(E218A)) strains in mid-log growth in Hv-Ca medium containing the indicated concentrations of Trp. **b**, **c**, Images of the strains were analysed for cell circularity (**b**) and the relative standard deviation (RSD) of curvature (**c**) of cell outlines. The wild-type (WT) control (H98+pTA962, 4 mM Trp) is shown in both **b** and **c**. At greater than 2 mM Trp, the circularity and curvature parameters did not increase further. However, during CetZ1(E218A) overproduction (4 mM Trp) cells appeared to show slightly larger highly distorted cells with more noticeable ruffling and protrusions (**a**, bottom right image). **d**, Coulter cell volume analysis

of the strains confirmed that CetZ1(E218A) overproduction (4 mM Trp) showed slightly larger cells, indicating delayed cell division—a possible effect of significant cell distortion. **e**, Western blot analysis of H98+pTA962-*cetZ1* (+CetZ1) and H98+pTA962-*cetZ1*(E218A) (+E218A) sampled during mid-log growth in Hv-Ca medium with the indicated concentrations of Trp. Ponceau S pre-staining of the membrane for total protein, as a loading reference, is shown below the CetZ1 western blot. The levels of CetZ1 increase in response to increasing concentrations of Trp in the medium. CetZ1(E218A) showed consistently higher levels compared to wild-type CetZ1, consistent with protofilament hyper-stability resulting in a slower rate of *in vivo* degradation of CetZ1(E218A) (see also Fig. 2a).



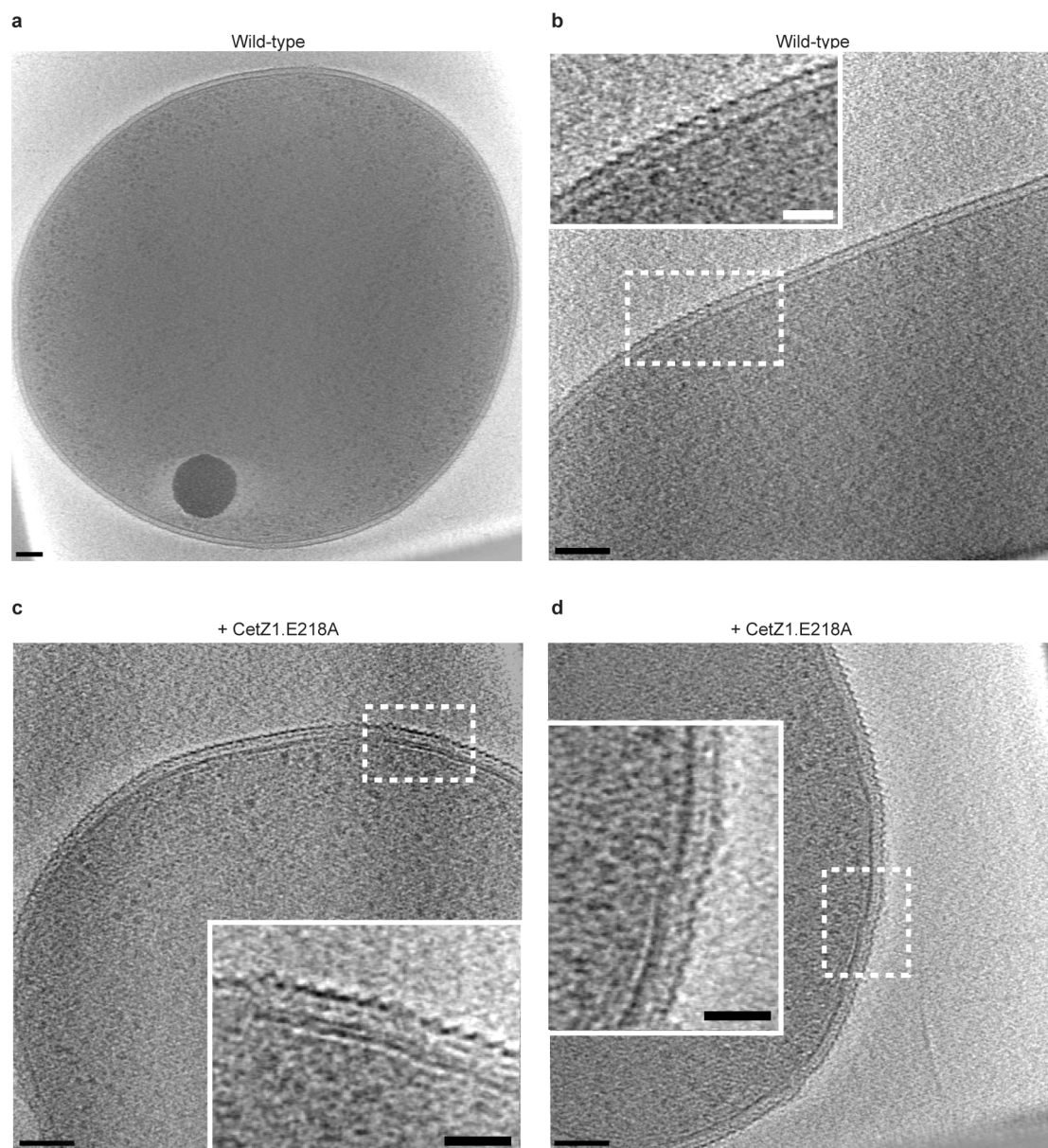
Extended Data Figure 5 | Localization of CetZ1-GFP and variants, CetZ1-GFP(A206K) and CetZ1-mCherry, in *H. volcanii*. CetZ1 tagged with the GFP variant used primarily in this study, smRS-GFP⁴⁰ was compared to CetZ1 tagged with the A206K mutant of smRS-GFP (which blocks GFP self-association^{51,52}) and the mCherry protein^{51,53} (which shows red fluorescence in

H. volcanii). Strains expressing the indicated CetZ1 fusion proteins in pTA962-based plasmids were sampled during mid-log phase growth in Hv-Ca + 0.2 mM Trp (log-phase cells) and from the leading edge of motile halos (0.2 mM Trp).



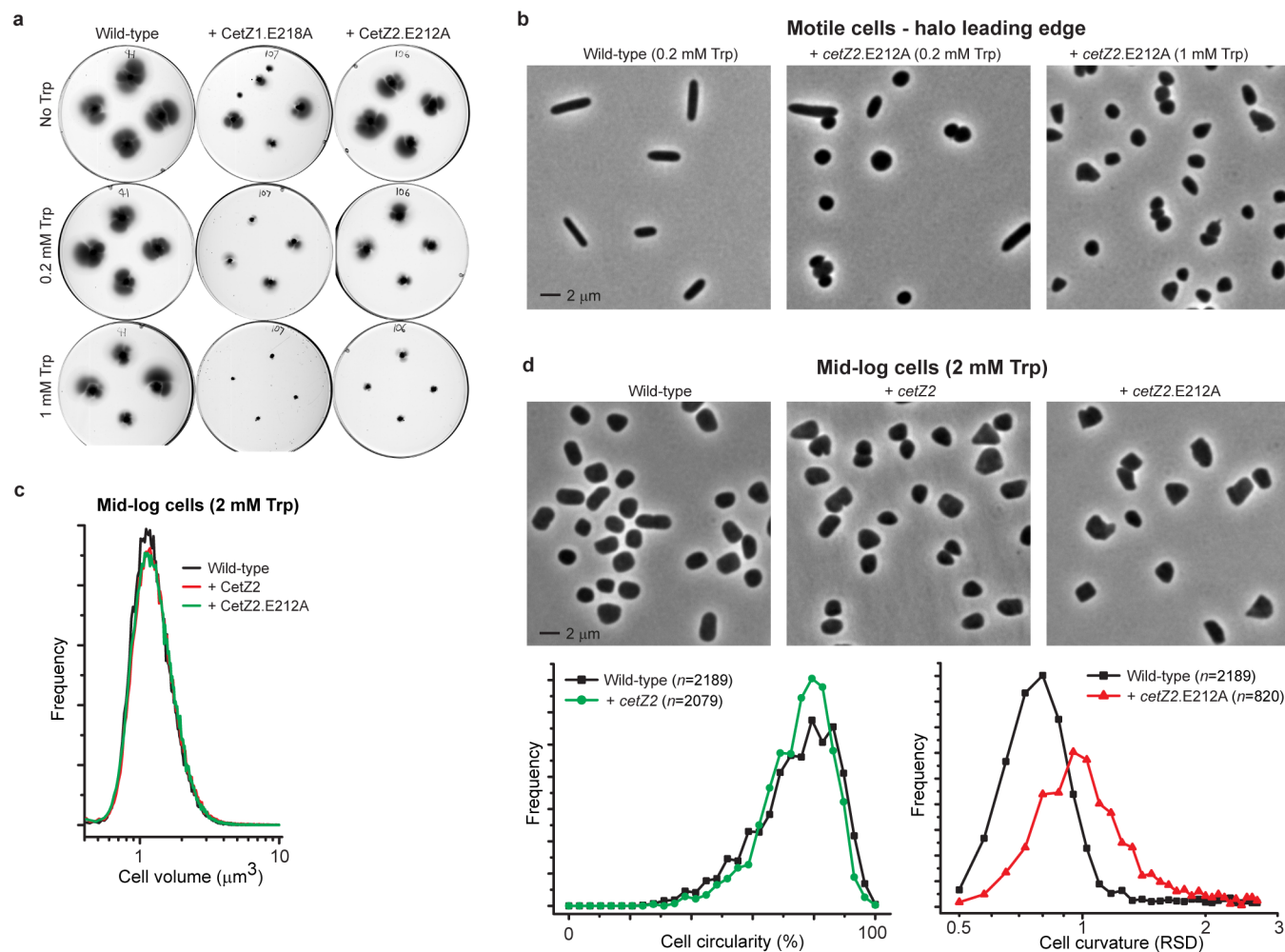
Extended Data Figure 6 | Expression of CetZ1-GFP and CetZ1(E218A)-GFP in *H. volcanii*. **a**, Phase-contrast (left) GFP-fluorescence (middle), and phase/GFP overlay (right) images of strains H98+pIDJL40 (GFP vector only), H98+pIDJL40-*cetZ1* (+CetZ1-GFP) and H98+pIDJL40-*cetZ1(E218A)* (+CetZ1(E218A)-GFP) sampled from mid-log growth in strong-overproduction conditions (Hv-Ca + 4 mM Trp). The GFP-vector-only fluorescence image is high-contrast, showing low-level background GFP fluorescence. The CetZ1-GFP and CetZ1(E218A)-GFP fluorescence images are presented with identical exposure and contrast settings, showing relatively higher background fluorescence for CetZ1-GFP and higher localized fluorescence for CetZ1(E218A)-GFP, indicating that assembly of the localized protein is stabilized by the E218A mutation (in the GTPase domain of CetZ). **b**, Cell circularity distributions for the indicated strains show no stimulation of rod shape during overproduction of either CetZ1-GFP or CetZ1(E218A)-GFP. **c**, Cell curvature distributions indicate that overproduction of CetZ1(E218A)-GFP results in blocky, stalked cells, as seen with untagged CetZ1(E218A) (Extended Data Fig. 4). **d**, Western blot analysis of CetZ1 in the indicated strains, showing increasing expression of CetZ1-GFP fusion proteins in response to Trp concentration. **e**, Motility assays for the indicated strains (H98 background, 0.2 mM Trp), showing significant

inhibition of motility by CetZ1(E218A); this may result from higher level production of this protein (see Extended Data Fig. 4e), or stronger inhibition of native CetZ1 function. **f**, Cell circularity analysis of cells at the leading edge of the halos in **e** suggests that strong inhibition of rod shape is needed to see an obvious reduction in motility. (Data for wild-type and +CetZ1(E218A) strains are reproduced from Fig. 2h.) **g**, Cell outlines resulting from the automated cell detection process in phase-contrast and the corresponding GFP-fluorescence images. Magnified images show the normal lines (3 pixels long) on the inside of the cell outline that were integrated to measure fluorescence near the envelope (the closely spaced normal lines give the appearance of a thick outline of the cell). Sections of high curvature (>2 standard deviations above the mean) are shown in green, whereas the remainder is shown in red. **h**, A plot of the curvature (red) and the GFP localization (green) of the magnified cell, measured anticlockwise from the arrowhead. In automated analysis of cell populations, the curvature of each whole cell was calculated as the relative standard deviation of this line. Analysis of the strain expressing CetZ1(E218A)-GFP (in Hv-Ca + 2 mM Trp) showed a 73% greater mean fluorescence in areas of high curvature versus the whole cell, 53% greater fluorescence for areas of high curvature versus the whole perimeter, and 10% greater fluorescence for the perimeter versus the whole cell.



Extended Data Figure 7 | Electron cryotomography of *H. volcanii*. Cells growing in Hv-YPC + 4 mM Trp were frozen in vitreous ice to maintain native structure, and then tomographic data were acquired with a 300 kV cryo-transmission electron microscope. A 10 nm section from the tomographic reconstruction of the whole cell is shown for each. **a**, H98 + pTA962 (wild-type) cell, showing cell envelope layers with some structural detail of the outer S-layer visible, a dense granule, particles consistent with ribosomes visible in the outer ~25% of the cytoplasmic region, and fibres that may be seen underlying

the cell envelope. **b**, Wild-type cell, with detail of the cell envelope region (inset as per Fig. 4f). **c**, Cells overproducing CetZ1(E218A) (H98 + pTA962-*cetZ1(E218A)*), show extensive sections of the envelope with underlying fibres (inset as per Fig. 4f). (Supplementary Video 5 shows a segment image series through a cell, showing the sheet-like structure of the extra layer.) **d**, Example showing subtle ruffling of the cell envelope associated with the additional envelope fibres in a cell producing CetZ1(E218A). Scale bars, 100 nm for all panels, and 50 nm for all insets.



Extended Data Figure 8 | Expression of CetZ2(E212A) affects *H. volcanii* cell shape and motility. The T7-loop mutation CetZ2(E212A) is equivalent to CetZ1(E218A). **a**, Expression of CetZ2(E212A) in motility assays, by addition of increasing concentrations of Trp, inhibits motility, but to a lesser degree than CetZ1(E218A). **b**, Cells were withdrawn from the leading edge of the indicated motile halos. Moderate expression of CetZ2(E212A) (0.2 mM Trp) partially inhibits rod development, and higher expression (1 mM Trp) strongly inhibits rod formation and produces cells with sharp corners and regions of high curvature, similar to cells expressing CetZ1(E218A) (Fig. 4). CetZ2(E218A) is therefore dominant-inhibitory to rod development. **c**, Coulter cell volume analysis of cells withdrawn from mid-log cultures of wild-type

(H98 + pTA962), CetZ2-overproducing (H98 + pTA962-*cetZ2*) and CetZ2(E212A)-overproducing (H98 + pTA962-*cetZ2*(E212A)) strains shows that overexpression of CetZ2 or CetZ2(E212A) does not affect cell division. **d**, Phase-contrast images of cells from the same cultures shown in **c**. Data obtained from automated analysis of images for each strain is shown in the plots in **d**. Unlike CetZ1 (Fig. 4), overexpression of CetZ2 in liquid culture did not stimulate rod morphology, indicating that CetZ2 is not a dominant driver of rod formation. However, analysis of cell curvature (bottom right) showed that expression of CetZ2(E212A) produced a high-curvature phenotype, similar to CetZ1(E218A) (Fig. 4), indicating that this mutant specifically interferes with shape regulation in *H. volcanii*.

Extended Data Table 1 | Crystallography data collection and refinement statistics

	CetZ2 – GTP γ S (PDB 4B45)	CetZ1 – GDP (PDB 4B46)	Mth CetZ – GDP (PDB 3ZID)
Data collection			
Space group	P 1	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	42.4, 43.2, 48.5	56.3, 62.2, 96.5	44.6, 107.0, 77.0
α , β , γ (°)	68.9, 75.3, 86.9	90.0, 90.0, 90.0	90.0, 91.42, 90.0
Resolution (Å)	40.3 – 2.0	48.7 – 1.9	50.0 – 2.0
<i>R</i> _{sym} or <i>R</i> _{merge} *	0.13 (0.94)	0.11 (0.58)	0.04 (0.66)
<i>I</i> / <i>σI</i> *†	8.1 (1.5)	10.0 (1.9)	11.4 (1.5)
Completeness (%)	97.9 (96.9)	99.8 (99.4)	92.9 (92.3)
Redundancy*	3.5 (3.6)	4.2 (3.3)	1.8 (1.8)
Refinement			
Resolution (Å)	2.1	1.9	2.0
No. reflections ‡	34258	50705	74471
<i>R</i> _{work} / <i>R</i> _{free} ‡§	0.18 (0.23)	0.17 (0.21)	0.18 (0.23)
No. atoms			
Protein	2502	2460	5531
Ligand/ion	32	28	56
Water	99	197	185
B-factors			
Protein	55.82	28.33	39.34
Ligand/ion	49.64	18.32	22.70
Water	53.64	36.37	41.89
R.M.S. deviations			
Bond lengths (Å)	0.007	0.006	0.008
Bond angles (°)	1.09	1.04	1.21

* The values recorded in parentheses refer to the highest resolution shell of the data set collected.

† Diffraction data was truncated at the resolution at which the half-set correlation coefficient fell below 0.7.

‡ The values quoted are those calculated and output by PHENIX.REFINE³⁹ during the final refinement run.

§ A set representing 5% of the reflections collected was randomly selected before refinement.

|| Mean atomic B-factors for each class of atom described in the final, deposited model.

Extended Data Table 2 | *H. volcanii* strains, plasmids and oligonucleotides

Strain	Genotype	Description*	Oligonucleotides used for construction (5' to 3')†	Source
DS70	(DS2) <i>ΔpHV2</i>	Wild-type <i>H. volcanii</i> cured of pHV2		T. Allers ⁴³
H53	(DS70) <i>ΔpyrE2 ΔtrpA</i>	Auxotroph		T. Allers ⁴⁰
H98	(DS70) <i>ΔpyrE2 ΔhdrB</i>	Auxotroph		T. Allers ⁴⁰
ID59	(H98) <i>ΔcetZ1</i>	Deletion of <i>cetZ1</i> (HVO_2204, Uniprot: D4GVD7)	1US-f: CCGGCCAAGCTTGCGAGTTCGTCTCCTTACGCA 1US-r: GCAGCACATCCCCCTTTCGCCAGATCTATGGGGACGTGTCGAACCAC 1DS-f: CCCCCAGATCTCTGGAGTCGCTTTTCTAAGCTGGC 1DS-r: CCCCCGGATCCACGTCTGCTCGGCTTGTGC	This study
ID58	(H98) <i>ΔcetZ2</i>	Deletion of <i>cetZ2</i> (HVO_0745, Uniprot: D4GTC1)	2US-f: CCGGCCAAGCTTGGGAGCCACTACCGTCACCAC 2US-r: GCAGCACATCCCCCTTTCGCCAGATCTCCTCACATGGGGACTGTTATCTC 2DS-f: CCCCCAGATCTAACCGACTCCTGTCTGTTTGC 2DS-r: CCCCCGGATCCGTCCGGTCTGCTCTCCAGTTTC	This study
ID60	(H98) <i>ΔcetZ3</i>	Deletion of <i>cetZ3</i> (HVO_1113, Uniprot D4GW48)	3US-f: CCGGCCAAGCTTGCTCAACGAGGGCGACACC 3US-r: GCAGCACATCCCGCTTTCGCCAGATCTGTGGTAGGGCTGTGCAAC 3DS-f: CCCCCAGATCTGCTGTTCTGACCGTCTTCTGTC 3DS-r: CCCCCGGATCCGAAGTCTCTCCGAGAGTCG	This study
ID61	(H98) <i>ΔcetZ4</i>	Deletion of <i>cetZ4</i> (HVO_A0035, Uniprot: D4GQ67)	4US-f: CCGGCCAAGCTTGCTGACTGAGTTCGTTCTGTTCTG 4US-r: GCAGCACATCCCCCTTTCGCCAGATCTATGGGACCCCTACCTACCC 4DS-f: CCCCCAGATCTTTTTAATGGATACCCCTGTCTCG 4DS-r: CCCCCGGATCCGAACACGCAACCGGCATTTC	This study
ID85	(H98) <i>ΔcetZ5</i>	Deletion of <i>cetZ5</i> (HVO_2013, Uniprot: D4GU26)	5US-f: CCGGCCAAGCTTGCTCTCCCGTTTTTCAGTTGC 5US-r: GCAGCACATCCCCCTTTCGCCAGATCTGTGGTACACGTCATCAGTGGTGAATC 5DS-f: CCCCCAGATCTCCGACCGCTCGACGATAG 5DS-r: CCCCCGGATCCGTTGCTACGCTGCTGCTT	This study
ID64	(H98) <i>ΔcetZ6</i>	Deletion of <i>cetZ6</i> (HVO_2068, Uniprot: D4GU81)	6US-f: CCGGCCAAGCTTGCTATCTCGATATCCCGTTCAA 6US-r: GCAGCACATCCCCCTTTCGCCAGATCTGTTGTCTGCTGCGACGTACC 6DS-f: CCCCCAGATCTAACGCAACGGGGTGCATAA 6DS-r: CCCCCGGATCCCTCAGCTCCGTCTCCGACTCA	This study
ID71	(H53) <i>ΔcetZ1</i>	As for ID59	As for ID59	This study

Plasmid	Description	Oligonucleotides used for construction (5' to 3')†	Source
pTA962	Expression vector. <i>tnaA</i> promoter. <i>pyrE2</i> and <i>hdrB</i> markers.		T. Allers ²¹
pIDJL40	<i>gfp</i> -fusion expression vector based on pTA962.‡	GFP-f: GGCCGGATCCATGAGTAAAGGAGAAGAACTTTTAC GFP-r: GGAAGAAATCGGGCCGCTTATTGTATAGTTTCATCCATGCCATG	This study
pIDJL117	<i>cetZ1-mCherry</i> expression plasmid based on pTA962.§	mCh-f: GGCCGGATCCGCTGGCTCCGCTGCTGGTTC mCh-r: GGAAGAAATCGGGCCGCTTACTTGTACAGCTCGTCCATGCC	This study
pTA962-ftsZ1.D250A	<i>ftsZ1.D250A</i> expression	ftsZ1-f: CCCCCGGGAATTCATATGGAAGTCTATCGTCGGCGACGC ftsZ1-r: CGCGGATCCCTACTCGACGTAGTCGATGCTTTCGAG F1.D250A-f: CCTCGACTTCGCGCGCTGAAGACCGTCATGG F1.D250A-r: CCATGACGGTCTTCACGCGCGCGAAGTCGAGG	This study
pIDJL40-ftsZ1	<i>ftsZ1-gfp</i> expression	ftsZ1-f: (shown above) ftsZ1(NS)-r : CGCGGATCCCTCGACGTAGTCGATGCTTTCGAG	This study
pTA962-cetZ1	<i>cetZ1</i> expression	cetZ1-f: CCCCCGGGAATTCATATGAAGCTCGCAATGATCGGATTCGGG cetZ1-r: CGCGGATCCCTTAGAAAAGCGACTCCAGTTCGTCTCTCG	This study
pTA962-cetZ1.E218A	<i>cetZ1.E218A</i> expression	cetZ1-f: (shown above) cetZ1-r: (shown above) C1.E218A-f: CGTCGCTGACTCCTCGGCGATTATCAACACGC C1.E218A-r: GCGTGTGATAATCGCCGAGGAGTCGACGACG	This study
pIDJL40-cetZ1	<i>cetZ1-gfp</i> expression	cetZ1-f: CCCCCGGGAATTCATATGAAGCTCGCAATGATCGGATTCGGG cetZ1(NS)-r: CGCGGATCCGAAAAGCGACTCCAGTTCGTCTCTCG	This study
pIDJL40-cetZ1.E218A	<i>cetZ1.E218A-gfp</i> expression	cetZ1-f and cetZ1(NS)-r (shown above) with pTA962- <i>cetZ1.E218A</i> PCR template	This study
pIDHV262	<i>cetZ1-gfp.A206K</i> expression	A206K-f: CCGTGCCACACAATCTAAGCTTTCGAAAGATCCC A206K-r: GGGATCTTTTCGAAAGCTTAGATTGTGTGGACAGG	This study
pTA962-cetZ2	<i>cetZ2</i> expression	cetZ2-f: CCCCCGGGAATTCATATGAAACCGTCTCTGATTGGTGTGGGG cetZ2-r: CGCGGATCCCTCACAGCAGGTCTGTCGAGGTCGTC	This study
pTA962-cetZ2.E212A	<i>cetZ2.E212A</i> expression	cetZ2-f: (shown above) cetZ2-r: (shown above) C2.E212A-f: CGTCGACACCTCCGCGGTCAATCAACACCTCC C2.E212A-r: GGAGGGTGTGTATGACCGCGGAGGTGTGCGAGC	This study

References 21, 40 and 43 are cited in this table.

* Locus identifier (HVO) and Uniprot database (<http://www.uniprot.org>) codes are given with the proposed *cetZ* gene names.

† PCR primers shown were used for amplification from DS70 or plasmid DNA as described in the Methods. Relevant restriction enzyme cutting sites are underlined. Complementary primers for generating point mutations were used with the ORF primers in overlap-extension reactions to generate the indicated mutant genes. US/DS-f/r – Upstream/Downstream-forward/reverse. (NS) – no stop codon.

‡ smRS-GFP was obtained from pJAM1020 (ref. 41) (from J. Maupin-Furrow). The internal NdeI site was removed by overlap extension PCR to incorporate a silent mutation (CATATG to CACATG), before cloning between BamHI and NotI of pTA962.

§ mCherry was obtained from pROD17 (from D. Sherratt). The internal PvuII site was removed by overlap-extension PCR to incorporate a silent mutation (CAGCTG to CAGCTC), and a mutation in the mCherry template was corrected during pIDJL117 construction (primers not shown).

Tetanus toxoid and CCL3 improve dendritic cell vaccines in mice and glioblastoma patients

Duane A. Mitchell^{1,2,3*}, Kristen A. Batich^{2,3*}, Michael D. Gunn^{4,5}, Min-Nung Huang⁵, Luis Sanchez-Perez², Smita K. Nair⁶, Kendra L. Congdon², Elizabeth A. Reap², Gary E. Archer^{1,2}, Annick Desjardins^{1,2}, Allan H. Friedman^{1,2}, Henry S. Friedman^{1,2}, James E. Herndon II⁷, April Coan⁷, Roger E. McLendon^{1,3}, David A. Reardon^{1,2}, James J. Vredenburgh^{1,2}, Darell D. Bigner^{1,2,3} & John H. Sampson^{1,2,3,5,8}

After stimulation, dendritic cells (DCs) mature and migrate to draining lymph nodes to induce immune responses¹. As such, autologous DCs generated *ex vivo* have been pulsed with tumour antigens and injected back into patients as immunotherapy. While DC vaccines have shown limited promise in the treatment of patients with advanced cancers^{2–4} including glioblastoma^{5–7}, the factors dictating DC vaccine efficacy remain poorly understood. Here we show that pre-conditioning the vaccine site with a potent recall antigen such as tetanus/diphtheria (Td) toxoid can significantly improve the lymph node homing and efficacy of tumour-antigen-specific DCs. To assess the effect of vaccine site pre-conditioning in humans, we randomized patients with glioblastoma to pre-conditioning with either mature DCs⁸ or Td unilaterally before bilateral vaccination with DCs pulsed with *Cytomegalovirus* phosphoprotein 65 (pp65) RNA. We and other laboratories have shown that pp65 is expressed in more than 90% of glioblastoma specimens but not in surrounding normal brain^{9–12}, providing an unparalleled opportunity to subvert this viral protein as a tumour-specific target. Patients given Td had enhanced DC migration bilaterally and significantly improved survival. In mice, Td pre-conditioning also enhanced bilateral DC migration and suppressed tumour growth in a manner dependent on the chemokine CCL3. Our clinical studies and corroborating investigations in mice suggest that pre-conditioning with a potent recall antigen may represent a viable strategy to improve anti-tumour immunotherapy.

To evaluate the influence of vaccine site pre-conditioning on DC migration clinically, we conducted a randomized and blinded clinical trial in

newly diagnosed glioblastoma (GBM) (Extended Data Fig. 1). A total of 13 patients consented to this trial but only 12 were randomized as 1 progressed before randomization (Extended Data Table 1). Patients were randomized to unilateral vaccine site pre-conditioning with unpulsed, autologous DCs⁸ or Td, on the basis of our hypothesis that it would induce inflammation at the vaccine site¹³. The accumulation of injected DCs in vaccine site-draining lymph nodes (VDLNs) was significantly greater in patients given Td (Fig. 1a). Moreover, Td-treated patients also showed a significant increase in both progression-free survival (Fig. 1b) and overall survival (Fig. 1c) compared to DC-treated patients. From the time of diagnosis, patients in the DC cohort had median progression-free and overall survivals of 10.8 and 18.5 months, respectively. Thus, the median progression-free and overall survivals for the DC cohort were consistent with patients treated with the standard of care¹⁴. Three censored patients from the Td cohort did not progress and were alive at the time of survival analysis (>36.6 months). Overall, these prognostic factors varied across both treatment groups as expected in a small clinical trial. However, there was no discernible trend across prognostic factors that would suggest that these factors alone account for the observed differences in survival between cohorts. Using both a previously published recursive partition analysis¹⁵ and the European Organization for Research and Treatment of Cancer (EORTC) nomogram¹⁶ for predicting outcome of patients with GBM, Td-treated patients exceeded expected survival times by a far greater degree than did DC-treated patients in both cases by nearly the same amount (Extended Data Table 1). The vaccine responses in long-term survivors varied in durability, but pp65-specific

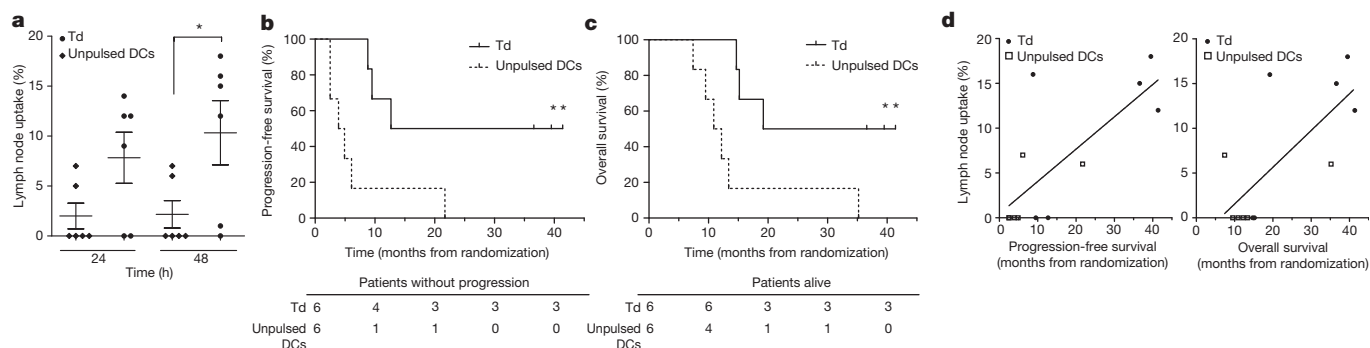


Figure 1 | Td pre-conditioning increases DC migration to VDLNs and is associated with improved clinical outcomes. **a**, DC migration in Td ($n = 6$) versus unpulsed DC ($n = 6$) patients (two sample t -test, $P = 0.049$). Mean \pm s.e.m., n values represent biological replicates of patient bilateral inguinal lymph nodes (iLNs). **b**, **c**, Patient progression-free survival (**b**) and overall survival (**c**) (log-rank test, $P = 0.013$). **d**, Hazard ratios (HRs): DC

migration efficiency from Td and DC cohorts showing the effect of a 1-unit increase in the percentage migration on progression-free (left) and overall (right) survival (Cox proportional hazards model, progression-free survival HR = 0.845 $P = 0.027$; overall survival HR = 0.820 $P = 0.023$). In **b** and **c**, $n = 3$ censored Td patients (no progressive disease at survival analysis).

¹Preston Robert Tisch Brain Tumor Center, Duke University Medical Center, Durham, North Carolina 27710, USA. ²Division of Neurosurgery, Department of Surgery, Duke University Medical Center, Durham, North Carolina 27710, USA. ³Department of Pathology, Duke University Medical Center, Durham, North Carolina 27710, USA. ⁴Division of Cardiology, Department of Medicine, Duke University Medical Center, Durham, North Carolina 27710, USA. ⁵Department of Immunology, Duke University Medical Center, Durham, North Carolina 27710, USA. ⁶Division of Surgical Sciences, Department of Surgery, Duke University Medical Center, Durham, North Carolina 27710, USA. ⁷Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, North Carolina 27710, USA. ⁸Department of Radiation Oncology, Duke University Medical Center, Durham, North Carolina 27710, USA.

*These authors contributed equally to this work.

immune responses were detectable for several months in all long-term survivors. An increase in pp65-specific interferon- γ spot-forming units from baseline did correlate with overall survival, and the two long-term survivors for which samples were available had the highest increases in pp65-specific immune responses after vaccination. In addition, we observed a notable association between DC migration to the VDLNs and progression-free and overall survival (Fig. 1d) in patients with GBM receiving pp65 RNA-pulsed DC vaccines.

To validate these clinical results and understand the mechanistic underpinnings, we performed analogous studies in a mouse model. Vaccine sites of Td-immune mice were pre-conditioned with Td and then received a bilateral vaccine of ovalbumin (OVA) RNA-pulsed DCs. In parallel to our clinical findings, Td-immune mice receiving Td pre-conditioning had a threefold increase in DCs within the afferent inguinal lymph nodes (Fig. 2a). This effect was attributable to Td-specific recall responses as mice not primed with Td (Td-naïve mice) did not display any increased DC migration to VDLNs (Fig. 2b). Vaccination and pre-conditioning with other CD4-dependent protein antigens also increased DC migration, suggesting that this may be a generalizable phenomenon (Extended Data Fig. 2). Subsequent studies performed in Td-treated mice demonstrated that only selective depletion of CD4⁺ T cells abrogated the increase in DC migration (Fig. 2c). The effect of enhanced migration was also transferable to naïve mice administered Td-activated CD4⁺ T cells (Fig. 2d).

In patients with GBM randomized to unilateral Td pre-conditioning (Fig. 2e) and in mice (Fig. 2f), we observed an increased uptake of ¹¹¹In-labelled DCs in bilateral lymph nodes, suggesting that Td pre-conditioning increased DC migration through systemic mediators. Subsequent experiments revealed that Td-activated CD4⁺ T cells administered systemically in naïve mice were also sufficient to increase bilateral DC migration (Extended Data Fig. 3). Vaccine site pre-conditioning with unpulsed DCs or TNF- α (ref. 8) only increased DC migration ipsilaterally (Extended Data Fig. 4a, b).

On the basis of our observations that Td recall responses could induce bilateral DC migration and that systemic administration of Td-activated

CD4⁺ T cells was sufficient to recapitulate the increased DC migration, we sought to examine the induction of CD4-dependent inflammatory mediators in the serum of patients and mice after a recall response with Td pre-conditioning. CCL3 was the only chemokine or cytokine to be increased in both patients and mice and had the greatest fold increase of all chemokines in the serum of both after Td pre-conditioning (Fig. 3a and Extended Data Fig. 5a–d).

To identify the site of the CCL3 production, we assayed the pre-conditioning sites in mice and found high concentrations of this chemokine only unilaterally at the site of Td pre-conditioning (Extended Data Fig. 6a). Subsequent experiments showed that CCL3 upregulation in the skin was dependent on the induction of the Td recall response (Extended Data Fig. 6b) and was significantly reduced by CD4⁺ T-cell depletion (Extended Data Fig. 6c). Induction of CCL3 by Td pre-conditioning remained increased over time compared to mice lacking Td recall responses (Extended Data Fig. 6d).

To evaluate the importance of CCL3 in mediating DC trafficking to VDLNs *in vivo*, we immunized and boosted *Ccl3*^{-/-} mice with Td as described above and found that the migration of injected DCs to VDLNs in *Ccl3*^{-/-} mice was significantly lower than in wild-type Td pre-conditioned mice (Fig. 3b, top). Endogenous migration experiments revealed that resident Langerhans cells also did not migrate as efficiently in *Ccl3*^{-/-} hosts after Td pre-conditioning (Extended Data Fig. 7). To address the possibility that Td priming may have failed in *Ccl3*^{-/-} mice owing to some earlier role of CCL3, we compared the CD4 T-cell-dependent immune responses to Td in both wild-type and *Ccl3*^{-/-} hosts and found no differences in the ability of these two strains to mount anti-tetanus immune responses after Td priming and boosting (Extended Data Fig. 8).

Our studies demonstrated that CD4⁺ T cells activated by the Td recall response were sufficient to induce increased DC migration in naïve mice (Fig. 2d). However, Td-activated CD4⁺ T cells could not rescue the limited DC migration in *Ccl3*^{-/-} hosts indicating that activated T cells were necessary, but not sufficient to increase DC migration (Fig. 3b, bottom). Exogenous administration of CCL3 could rescue the limited

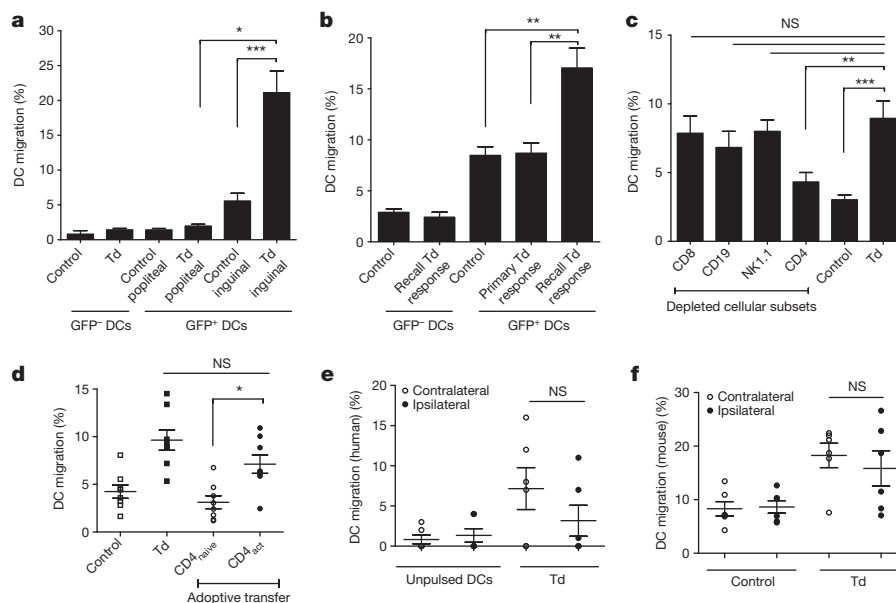


Figure 2 | Td recall response activates CD4⁺ T cells to increase DC migration to VDLNs. **a**, Control inguinal ($n = 5$ mice) versus Td inguinal ($n = 5$ mice); two sample *t*-test, $P = 0.0001$; Td popliteal ($n = 5$ mice) versus Td inguinal, paired *t*-test, $P = 0.014$. **b**, Mice primed and boosted with saline (primary Td, $n = 6$) or Td (control and recall Td, $n = 6$) with Td (primary and recall Td) or saline (control) pre-conditioning; one-way analysis of variance (ANOVA), $P = 0.004$; post-hoc Tukey *t*-test, control versus recall Td, $P = 0.006$; primary Td versus recall Td, $P = 0.011$. **c**, DC migration in depleted Td-immunized mice ($n = 5$); one-way ANOVA, $P < 0.0001$; post-hoc Tukey

t-test, Td versus CD4, $P = 0.005$; Td versus CD8, CD19 or NK1.1, $P > 0.05$. **d**, DC migration after CD4⁺ transfer ($n = 4$ mice); one-way ANOVA, $P < 0.0001$; post-hoc Tukey *t*-test, Td-activated CD4⁺ T cells (CD4_{act}) versus CD4_{naive}, $P < 0.05$; control versus CD4_{naive}, $P > 0.05$; Td versus CD4_{act}, $P > 0.05$. **e**, Patient (**e**) and mouse (**f**) iLN ipsilateral ($n = 6$) and contralateral ($n = 6$) to pre-conditioning; paired *t*-test, $P = 0.28$ (**e**) and $P = 0.37$ (**f**). **a–d** and **f** are representative of four experiments; mean \pm s.e.m. Data in **a–d** denote biological replicates of individual right and left iLN or lymph nodes ipsilateral to Td/saline for popliteal and GFP⁺ groups.

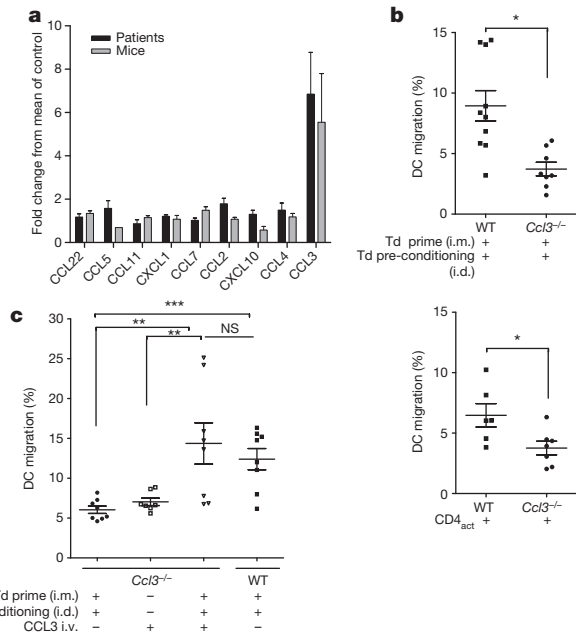


Figure 3 | Td recall responses and induced CCL3 cooperate to facilitate DC migration to VDLNs. **a**, Serum CCL3 fold increase over non-Td cohorts (patient, $n = 6$; mouse, $n = 8$ biological replicates); signed-rank test, $P = 0.031$ and $P = 0.039$. **b**, Top, DC migration in wild-type (WT) ($n = 5$) versus $Ccl3^{-/-}$ ($n = 4$) mice; two sample t -test, $P = 0.023$. Bottom, Td-activated $CD4^{+}$ transfer in $Ccl3^{-/-}$ hosts ($n = 4$ mice); two sample t -test, $P = 0.029$. **c**, CCL3 and Td recall responses rescue migration ($n = 4$ mice); one-way ANOVA, $P < 0.0001$; post-hoc Tukey t -test, $Ccl3^{-/-}$ Td plus Td i.d. and CCL3 i.v. versus $Ccl3^{-/-}$ Td plus Td i.d. and versus $Ccl3^{-/-}$ plus CCL3 i.v., $P = 0.007$ and $P = 0.001$, respectively. Data in **b** and **c** denote biological replicates of individual right and left ILNs. Representative of three experiments; mean \pm s.e.m. i.d., intradermal; i.m., intramuscular; i.v., intravenous.

DC migration in $Ccl3^{-/-}$ mice, but only when the Td recall responses were induced (Fig. 3c), indicating that the ability of Td pre-conditioning to increase DC migration to VDLNs was dependent on both $CD4^{+}$ recall responses and host-derived CCL3.

Alterations in CCL21 expression along the lymphatic endothelium in the skin take place in the context of inflammation and have been associated with increased DC migration⁶. After Td pre-conditioning, higher levels of CCL21 in collected skin sites were detected in Td-treated wild-type mice compared to $Ccl3^{-/-}$ mice, and exogenous administration of CCL3 rescued the diminished levels of CCL21 at the vaccine sites of $Ccl3^{-/-}$ mice only in the context of Td recall responses (Extended Data Fig. 9a), which paralleled the requirement of host CCL3 and Td recall responses for increased DC migration. Furthermore, Td pre-conditioning resulted in an increase in CCL21 within inguinal lymph nodes both ipsilateral and contralateral to the side of Td pre-conditioning (Extended Data Fig. 9b). Although lymph nodes contain high basal amounts of CCL21, we found CCL21 levels in bilateral inguinal lymph nodes of Td-treated $Ccl3^{-/-}$ mice to be lower than that of wild-type mice. Meanwhile, reconstitution of CCL3 back into $Ccl3^{-/-}$ mice in conjunction with the induction of Td recall responses markedly increased the expression of CCL21 in VDLNs (Extended Data Fig. 9c).

The apparent increase in progression-free and overall survival for Td treated-patients prompted us to determine whether Td pre-conditioning could inhibit tumour growth in a mouse subcutaneous tumour model in which B16F10 melanoma cells expressed OVA (B16-OVA). Pairwise comparisons revealed that tumour growth in Td plus OVA-DC mice was significantly delayed (Fig. 4a) in an antigen-dependent manner (Fig. 4b). Td pre-conditioning in $Ccl3^{-/-}$ mice, as expected, was unable to inhibit tumour growth (Fig. 4c). Similarly, *plt* (paucity of lymph node T cell) mice, which lack expression of CCL21 in the lymph nodes¹⁷,

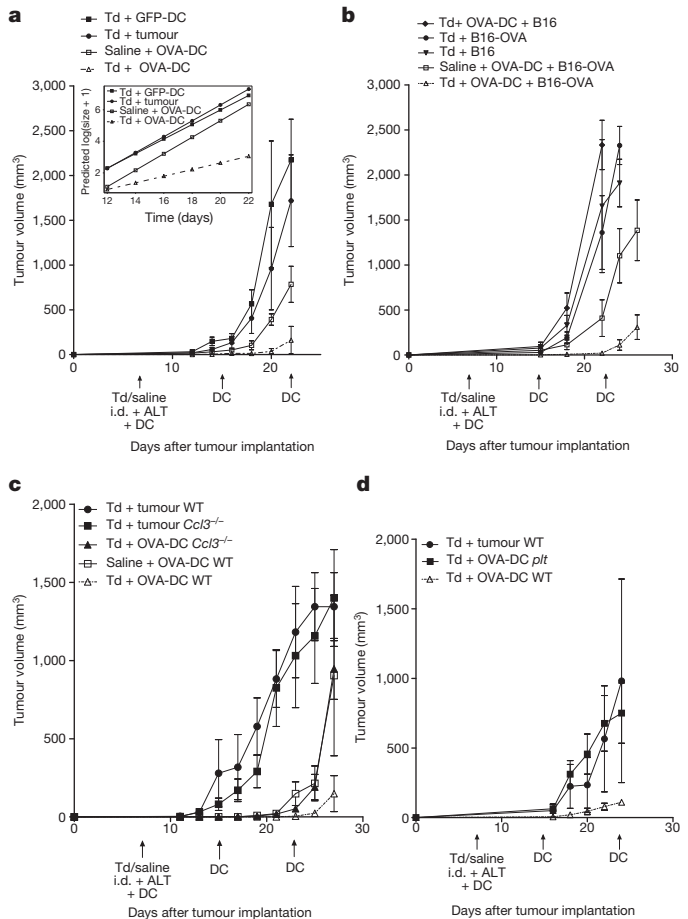


Figure 4 | Td pre-conditioning improves responses in tumour-bearing mice. **a**, Inset, transformed growth curves, mixed linear effects model. Pairwise comparisons of regression line slopes (F -test, $P < 0.0001$). Day 22 volume (Td plus OVA-DC versus Td plus GFP-DC, two sample t -test, $P = 0.002$; $n = 7$). **b**, Antigen-specific responses with Td pre-conditioning. Day 15 volume (all groups, F -test, $P = 0.016$; pairwise Tukey t -tests, Td plus OVA-DC and B16-OVA ($n = 7$) versus Td plus OVA-DC and B16 ($n = 5$), $P = 0.0004$; Td plus OVA-DC and B16-OVA versus saline plus OVA-DC and B16-OVA ($n = 6$), $P = 0.0002$). Day 22 volume (Td plus OVA-DC and B16-OVA versus Td plus OVA-DC and B16, two sample t -test, $P < 0.0001$). **c**, Tumour growth in $Ccl3^{-/-}$ mice. Day 11 volume (all groups, F -test, $P = 0.005$). Day 27 volume (Td plus OVA-DC WT versus Td plus OVA-DC $Ccl3^{-/-}$, two sample t -test, $P = 0.042$; $n = 8$). **d**, Anti-tumour responses in *plt* mice. Day 16 volume (all groups, F -test, $P = 0.004$). Day 24 volume (Td plus OVA-DC *plt* ($n = 7$) versus Td plus OVA-DC WT ($n = 6$), two sample t -test, $P < 0.05$). Data in **a–d** are representative of three experiments; mean \pm s.e.m. ALT, autologous lymphocyte transfer.

also failed to suppress tumour growth (Fig. 4d), indicating that Td-enhancement of OVA-DC vaccines also required intact CCL21 expression in the host.

Results from our trial seem to demonstrate that the modulation of CMV-specific DCs with Td pre-conditioning increases their migratory capacity and may improve clinical outcomes in patients with GBM. Corroborating studies in mice support these claims and underline CCL3 as a novel and important mediator of increased DC migration to VDLNs, in addition to its described roles in DC precursor mobilization to peripheral sites of inflammation^{18,19} and in guiding naive $CD8^{+}$ localization to sites of DC- $CD4^{+}$ interactions in lymph nodes²⁰. Our preclinical findings suggest that increased DC migration was only observed in the context of host CCL3 and Td recall, which may have a role in DC homing to local lymph nodes. However, owing to the already high basal amount

of CCL21 chemokine in draining lymph nodes, it is possible that additional factors such as interactions between host CCL3 and CD4⁺ recall T cells are the predominant contributors to the increased DC accumulation in VDLNs. Our findings also suggest that DC migration should be further investigated as a predictive biomarker for immunotherapy studies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 6 December 2013; accepted 13 February 2015.

Published online 11 March 2015.

- Steinman, R. M. & Banchereau, J. Taking dendritic cells into medicine. *Nature* **449**, 419–426 (2007).
- Palucka, A. K. *et al.* Spontaneous proliferation and type 2 cytokine secretion by CD4⁺T cells in patients with metastatic melanoma vaccinated with antigen-pulsed dendritic cells. *J. Clin. Immunol.* **25**, 288–295 (2005).
- Palucka, A. K. *et al.* Single injection of CD34⁺ progenitor-derived dendritic cell vaccine can lead to induction of T-cell immunity in patients with stage IV melanoma. *J. Immunother.* **26**, 432–439 (2003).
- Palucka, A. K. *et al.* Dendritic cells loaded with killed allogeneic melanoma cells can induce objective clinical responses and MART-1 specific CD8⁺ T-cell immunity. *J. Immunother.* **29**, 545–557 (2006).
- Palucka, K. & Banchereau, J. Cancer immunotherapy via dendritic cells. *Nature Rev. Cancer* **12**, 265–277 (2012).
- Liau, L. M. *et al.* Dendritic cell vaccination in glioblastoma patients induces systemic and intracranial T-cell responses modulated by the local central nervous system tumor microenvironment. *Clin. Cancer Res.* **11**, 5515–5525 (2005).
- Yu, J. S. *et al.* Vaccination with tumor lysate-pulsed dendritic cells elicits antigen-specific, cytotoxic T-cells in patients with malignant glioma. *Cancer Res.* **64**, 4973–4979 (2004).
- Martin-Fontecha, A. *et al.* Regulation of dendritic cell migration to the draining lymph node: impact on T lymphocyte traffic and priming. *J. Exp. Med.* **198**, 615–621 (2003).
- Dziurzynski, K. *et al.* Consensus on the role of human cytomegalovirus in glioblastoma. *Neuro Oncol.* **14**, 246–255 (2012).
- Ranganathan, P., Clark, P. A., Kuo, J. S., Salamat, M. S. & Kalejta, R. F. Significant association of multiple human cytomegalovirus genomic loci with glioblastoma multiforme samples. *J. Virol.* **86**, 854–864 (2012).
- Mitchell, D. A. *et al.* Sensitive detection of human cytomegalovirus in tumors and peripheral blood of patients diagnosed with glioblastoma. *Neuro Oncol.* **10**, 10–18 (2008).
- Cobbs, C. S. *et al.* Human cytomegalovirus infection and expression in human malignant glioma. *Cancer Res.* **62**, 3347–3350 (2002).
- Myers, M. G., Beckman, C. W., Vording, R. A. & Hankins, W. A. Primary immunization with tetanus and diphtheria toxoids. Reaction rates and immunogenicity in older children and adults. *J. Am. Med. Assoc.* **248**, 2478–2480 (1982).
- Stupp, R. *et al.* Radiotherapy plus Concomitant and Adjuvant Temozolamide for Glioblastoma. *N. Engl. J. Med.* **352**, 987–996 (2005).
- Curran, W. J. Jr. *et al.* Recursive partitioning analysis of prognostic factors in three Radiation Therapy Oncology Group malignant glioma trials. *J. Natl Cancer Inst.* **85**, 704–710 (1993).
- Gorlia, T. *et al.* Nomograms for predicting survival of patients with newly diagnosed glioblastoma: prognostic factor analysis of EORTC and NCIC trial 26981–22981/CE3. *Lancet Oncol.* **9**, 29–38 (2008).
- Nakano, H. & Gunn, M. D. Gene duplications at the chemokine locus on mouse chromosome 4: multiple strain-specific haplotypes and the deletion of secondary lymphoid-organ chemokine and EBI-1 ligand chemokine genes in the plt mutation. *J. Immunol.* **166**, 361–369 (2001).
- Zhang, Y. *et al.* Mobilization of dendritic cell precursors into the circulation by administration of MIP-1 α in mice. *J. Natl Cancer Inst.* **96**, 201–209 (2004).
- He, S. *et al.* MIP-3 α and MIP-1 α rapidly mobilize dendritic cell precursors into the peripheral blood. *J. Leukoc. Biol.* **84**, 1549–1556 (2008).
- Castellino, F. *et al.* Chemokines enhance immunity by guiding naive CD8⁺ T cells to sites of CD4⁺ T cell-dendritic cell interaction. *Nature* **440**, 890–895 (2006).

Acknowledgements The authors thank the staff who supported this study, including R. Schmittling, P. Norberg, W. Xie, P. Healy, D. Lally-Goss, S. McGehee-Norman, B. Perry, S. Snipes and R. Edward Coleman. This work was supported by grants from the National Institutes of Health (NIH) National Institute of Neurological Disorders and Stroke Specialized Program of Research Excellence in brain cancer (P50-CA108786, D.D.B. and J.H.S.) and SRC on Primary and Metastatic Tumors of the CNS (P50-NS20023, D.D.B. and J.H.S.) as well as NIH R01 (R01-CA177476-01, J.H.S.; R01-NS067037, R01-CA134844, D.A.M.) and P01 (P01-CA154291-01A1, D.D.B. and J.H.S.) funding sources. Additional support is from the National Brain Tumor Society (D.A.M. and J.H.S.), the American Brain Tumor Association (D.A.M. and J.H.S.), Accelerate Brain Cancer Cure Foundation Young Investigator's Award (D.A.M.), The Kinetics Foundation, (J.H.S.) Ben and Catherine Ivy Foundation (J.H.S.), and in part by Duke University's Clinical & Translational Science Awards grant 1UL2 RR024128-01 from the National Institutes of Health National Center for Research Resources.

Author Contributions D.A.M., G.E.A. and J.H.S. jointly conceived and implemented the clinical study; D.A.M. and K.A.B. jointly designed early DC migration studies in mice. K.A.B. conceived and designed the remainder of the mouse research; A.D., A.H.F., H.S.F., R.E.M., D.A.R., J.J.V., D.D.B. and J.H.S. were responsible for provision of clinical study resources, materials and patient access. S.K.N., E.A.R. and G.E.A. prepared human samples and conducted human *in vitro* experiments; K.A.B. performed all preclinical experiments and patient analyses; M.-N.H. provided additional support for preclinical experiments. D.A.M., K.A.B., M.D.G., K.L.C., E.A.R., G.E.A. and J.H.S. performed data analysis and interpretation; J.E.H. and A.C. provided statistical support for design and analysis of human and mouse studies. D.A.M., D.D.B. and J.H.S. contributed laboratory reagents and tools; and D.A.M., K.A.B., M.D.G., K.L.C. and J.H.S. wrote the paper. All authors gave their final approval to the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare competing financial interests: details are available in the online version of the paper. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to J.H.S. (john.sampson@duke.edu) or D.A.M. (duane.mitchell@neurosurgery.ufl.edu).

METHODS

Patient selection, demographics and clinical protocol. The clinical protocol and informed consent were approved by the US Food and Drug Administration and Institutional Review Board at Duke University. Adults with a newly diagnosed World Health Organization (WHO) grade IV GBM, who had a gross total resection and residual radiographic contrast enhancement on post-resection magnetic resonance imaging (MRI) not exceeding 1 cm in diameter in two perpendicular axial planes, and a Karnofsky performance scale score of ≥ 80 , were eligible for the clinical study (FDA - IND-BB-12839, Duke IRB Pro00003877, NCT00639639). Histopathology of all specimens was initially read as GBM, but this diagnosis was re-confirmed by a second board-certified neuropathologist. Histological diagnosis included immunohistochemistry for MGMT protein expression. Benign endothelial cells staining positive for MGMT served as the internal control²¹. MGMT promoter methylation was performed by PCR. On the basis of published reports showing high expression of CMV viral proteins in $>90\%$ of GBM tumours^{9–12}, we elected not to include pp65 staining of tumour tissue as an eligibility criterion for this trial. All 13 patients on study received a gross total resection defined as $>90\%$ with residual contrast enhancement of $<1\text{ cm}^2$, and steroid doses could not exceed 2 mg day^{-1} of dexamethasone. No patients received intensity-modulated radiation therapy or had 5-aminolevulinic acid dye used during resection. Thereafter, all patients completed a 6-week course of conformal external beam radiotherapy to a dose of 60 Gy with concurrent temozolomide at a targeted daily dose of $75\text{ mg m}^{-2}\text{ day}^{-1}$. After completion of standard therapy, all patients underwent an MRI for evidence of progressive disease. Those with evidence of progressive disease or required steroid therapy in excess of physiological levels at the time of vaccination were replaced. A total of 13 patients were enrolled and randomized before the first cycle of standard-of-care 5-day TMZ ($200\text{ mg m}^{-2}\text{ day}^{-1}$), but one progressed before randomization. For each vaccine, 2×10^7 mature pp65 RNA-pulsed DCs in 0.4 ml of saline were given intradermally in the groin. The first vaccination occurred on day 21 ± 2 of TMZ cycle 1. Although some patients ($n = 5$) were also randomized to receive an autologous lymphocyte transfer, those patients did not show a significant improvement in progression-free or overall survival. Patients given autologous lymphocytes were additionally administered 3×10^7 cells kg^{-1} intravenously with acetaminophen ($650\text{ mg per os (po)}$) and Benadryl ($25\text{--}50\text{ mg po}$) given 30–60 min before infusion. The first three DC vaccines were given bi-weekly, and, at vaccine 4, patients were randomized to Td or unpulsed autologous DCs and received ¹¹¹In-labelled DCs for migration studies. Vaccine 4 and additional monthly vaccines until tumour progression occurred on day 21 ± 2 of successive TMZ cycles. A minimum of six cycles of adjuvant TMZ were required as per standard-of-care and continuation was at the discretion of the treating neurooncologist. Patients were monitored for treatment-related toxicity, and none of the patients experienced any vaccine or Td-related adverse events.

Human autologous DC generation for vaccination and production of pp65-LAMP/A64 mRNA. DCs were generated using the method described previously²², and after collection the cells were frozen and assessed for contamination and lineage purity as previously published²³. The 1.932-kilobase (kb) pp65 full-length cDNA insert was obtained from B. Britt and RNA was generated and transfected as previously reported²².

Human DC migration studies. DC migration studies were done at the fourth vaccination. Patients were randomized by side to have one inguinal vaccination site pre-treated with either 1×10^6 unpulsed DCs or Td toxoid (1 flocculation unit (Lf)). Saline was administered on the contralateral side. Vaccination site pretreatment was done 6–24 h before DC vaccination. DCs were labelled with $10\text{ }\mu\text{Ci per }1 \times 10^7$ DC with ¹¹¹In (GE Healthcare) and divided equally in the two sites. Gamma camera images (GE Infinia Hawkeye) were taken immediately after injection and at 24 and 48 h after injection to compare ¹¹¹In-labelled DC migration from the inguinal injection sites to the inguinal lymph nodes.

Progression-free and overall survival. The more recent response evaluation criteria in solid tumours (RECIST criteria) judge progression by measuring the longest one-dimensional diameter and determine progression by a 20% increase in this diameter²⁴. Once progression is detected on MRI, other imaging modalities such as positron emission tomography and a stereotactic brain biopsy of the enhancing region are incorporated to aid in determining progression. A stereotactic brain biopsy or resection demonstrating recurrence defines clinical progression. Progression-free survival was defined as the time until radiographic or clinical progression and was censored at the last follow-up if the patient remained alive without disease progression. Overall survival was defined as the time until death and was censored at the last follow-up if the patient remained alive at the time of analysis. Progression-free and overall survival for all patients were calculated from both the time of surgery and from randomization to vaccine site pre-conditioning.

Mice. All animal experiments were performed according to Duke University Institutional Animal Care and Use Committee-approved protocols. Female C57BL/6 wild-type, OT-I transgenic mice, *Ccl3*^{−/−}, and red fluorescent protein (RFP) and

green fluorescent protein (GFP) transgenic mice (ubiquitin promoter) were obtained from the Jackson Laboratory and were bred under pathogen-free conditions at Duke University Medical Center. The *plt* strain was provided by M.D.G. and maintained at Duke University Medical Center. All mice were bred under pathogen-free conditions at Duke University Medical Center.

Generation of mouse bone marrow-derived DCs, electroporation and phenotyping. Bone-marrow-derived DCs were generated from 6–8-week-old female C57BL/6 wild-type, RFP⁺ or GFP⁺ transgenic mice and pulsed with OVA RNA as previously described²⁵. For phenotyping, anti-mouse phycoerythrin (PE)-conjugated CD11c (HL3), CD80 (16-10A1), CD86 (GL1), Ly-6G (1A8), MHC class II (I-A^b; AF6-120.1) and isotype controls (IgG1; G235-2356, IgG2a,κ; R35-95) were from BD Pharmingen. Cells were washed, resuspended in PBS and 2% FBS, incubated at 4 °C for 30 min, and washed again before use.

Vaccine site pre-conditioning and DC vaccination in mice. For Td immunization, female 6–8-week-old C57BL/6 mice received a primary intramuscular vaccine of Td toxoid (Sanofi Aventis; Decavac; 1 Lf, 100 μl) administered bilaterally into the quadriceps muscle (50 μl per leg). An intramuscular booster (0.5 Lf, 50 μl) was administered two weeks later. Vaccine site pre-conditioning with saline or Td toxoid (0.5 Lf) was given intradermally (i.d.) 2 weeks after the booster and randomized to the right or left groin site. Mouse IgG antibody responses to Td were measured by ELISA (Xpress Bio). Serum from immunized mice was collected 2 weeks after each immunization before the next booster vaccine. DCs were resuspended at 1×10^6 per 100 μl PBS (Gibco) and administered i.d. on both sides 0.8 cm from the groin crease 24 h after i.d. pre-conditioning. DCs injected in the groin ipsilateral to the Td pre-conditioning side were directly injected i.d. within the erythematous nodule produced by Td pre-conditioning. For recall response experiments using other protein antigen formulations, female 6–8-week-old C57BL/6 mice received a primary intramuscular vaccine of Pevnar 13 (Pfizer, Pneumococcal 13-valent conjugate vaccine, 1.32 μg , 100 μl) and Pedvax HIB (Merck, Haemophilus b conjugate vaccine, 1.5 μg , 100 μl) administered bilaterally into the quadriceps muscle (50 μl per leg). Vaccine site pre-conditioning with saline or the protein antigen (50 μl) was given i.d. 2 weeks later and randomized to the right or left groin site. DC vaccines were given 24 h later, and migration to lymph nodes was assessed 48 h later. As with Td pre-conditioning, DCs injected in the groin ipsilateral to the pre-conditioning side were directly injected i.d. within the erythematous nodule produced by those formulations. For comparisons of other pre-conditioning agents, female 6–8-week-old C57BL/6 mice received a unilateral dose of unpulsed, mature DCs (1×10^6 in 50 μl) or TNF- α (30 ng) administered i.d. at the groin site 24 h before DC vaccination. On the basis of the previous work using these pre-conditioning regimens, DC migration to bilateral inguinal lymph nodes was assessed 24 h later. For all other migration experiments, popliteal and inguinal lymph nodes were collected 48 h after DC vaccination and digested for flow cytometry. The percentage of migrating DCs was enumerated by gating on fluorescent DCs in wild-type VDLNs. DCs from wild-type (GFP[−] and RFP[−]) mice as negative controls before gating on fluorescent DCs within VDLNs to account for background autofluorescent cells that may have appeared in the GFP channel. For *in vivo* DC migration, a sample size (three per group) was based on empirical evidence from previously published reports as the size necessary for adequate statistical analysis of lymph nodes sampled²⁶.

Depletion, adoptive transfer and CCL3 reconstitution. Female 6–8-week-old C57BL/6 mice were initially depleted of cellular subsets once daily (200 μg per mouse intraperitoneally) for 3 days before the first Td intramuscular immunization. Anti-mouse CD4 (GK1.5) and anti-CD8 (2.43) antibodies were purchased from American Type Culture Collection. Anti-mouse NK1.1 (PK136) and anti-CD19 (2D5) and control isotype depleting antibodies (IgG2a (2A3) and IgG2b (LTF-2)) were from BioXCell. Maintenance doses of depletion antibodies were administered at 3-day intervals (200 μg intraperitoneally) until vaccine site pre-conditioning with Td 2 weeks later. For adoptive transfer experiments, Td-activated CD4⁺ T cells (CD4^{act}) were induced in donor female 6–8-week-old C57BL/6 mice. Mice were primed (1 Lf, 100 μl) and boosted (0.5 Lf, 50 μl) intramuscular with Td 2 weeks apart. Three days after the i.d. Td pre-conditioning, donor inguinal lymph nodes, skin injection sites, and spleens were collected and processed for negative isolation of CD4⁺ T cells (Miltenyi Biotec). Complementary sites from naive mice were collected simultaneously and processed for negative isolation of CD4⁺ T cells (CD4^{naive}). A final dose of 6×10^6 CD4⁺ T cells were administered intravenously into recipient mice two days before i.d. vaccination with RFP⁺ DCs. For CCL3 reconstitution in *Ccl3*^{−/−} hosts, recombinant mouse CCL3 (R&D Systems) was administered intravenously into the tail vein (10 μg per mouse) 12 h before vaccination with RFP⁺ DCs. *Ccl3*^{−/−} mice that were Td-immune were given recombinant CCL3 12 h after Td pre-conditioning at the vaccine site.

Tumour implantation experiments. For tumour implantation experiments, B16F10-OVA cells were grown as previously published²⁷ and injected subcutaneously at a concentration of 2×10^5 cells in 200 μl of PBS in the flank of C57BL/6 mice 8 days before vaccine site pre-conditioning, the first intradermal vaccine of OVA RNA-pulsed

DCs, and autologous lymphocyte transfer (1:1 infusion of naive:OT-I OVA-specific T cells). Randomization of mice occurred after tumour inoculation before vaccine site pre-conditioning and the first DC vaccine first by compilation and then by random sorting into various treatment cages. Mice received two additional weekly vaccines of RNA-pulsed DCs on days 15 and 22. Ten days after tumour implantation, flank sites were monitored daily for tumour growth, and tumour size was measured every 2 days. Tumour volume (millimetres cubed) was calculated by the formula (length \times width² \times 0.52) in a perpendicular fashion. Mice were euthanized when ulceration occurred or when the tumour reached either 2 cm in any direction or 2,000 mm³. Analysis of tumour growth focused on follow-up assessments before considerable dropout occurred. A logarithmic transformation yielded a linear relationship between tumour volume and time for all curves. A mixed effects linear model that accounted for correlation of measurements within a mouse was used to examine the relationship between time and log [tumour volume + 1]. No blinding was done for these animal studies.

Mouse tumour cell lines. The B16F10-OVA tumour cell line was a gift from R. Vile^{27,28}. The B16F10 cell line was provided by I. Fidler²⁹. Cell lines were tested for mycoplasma before use.

Mouse lymph node digestion and quantification of fluorescent and endogenous DCs. Lymph nodes were placed in 6-well culture plates containing 1 ml HBSS with Ca²⁺/Mg²⁺ (Gibco), digested for 35 min at 37 °C with collagenase A (1 mg ml⁻¹; Roche) and DNaseI (0.2 mg ml⁻¹; Sigma-Aldrich) and 20 mM EDTA (Invitrogen) was added for 5 min at room temperature to stop the reaction²⁶. Single-cell suspensions were prepared, cells were centrifuged (500g for 5 min) and resuspended in PBS with 2% FBS and stained with mouse allophycocyanin (APC)-conjugated CD11c (BD Pharmingen; HL3). For quantification of RFP⁺ or GFP⁺ counts in individual lymph nodes, samples were resuspended at an equal volume and 50 µl of counting beads (Invitrogen; 50,000 beads) were added to each sample. Cells were gated first on mouse CD11c⁺ cells and then RFP⁺ or GFP⁺ cells, and absolute cell counts/lymph nodes were quantified using the following equation: RFP⁺ or GFP⁺ events \times 50,000 beads/number of bead events. For endogenous DC migration experiments, cells were surface stained in PBS with 3% FBS, 10 mM EDTA, 5% normal mouse serum, 5% normal rat serum and 1% Fc Block (eBioscience; clone 93) and then intracellularly stained with anti-CD207 according to the manufacturer's protocol (BD Cytofix/Cytoperm Kit). The cells were analysed by BD LSRII flow cytometer with FlowJo software (Tree Star). FITC-conjugated anti-B220 (RA-3-6B2), Alexa Fluor 700-conjugated anti-Ly-6G (1A8), APC-Cy7-conjugated anti-CD11b (M1/70), V450-conjugated anti-Ly-6C (AL-21) are from BD Pharmingen. PE-conjugated anti-CD207 (eBioL31), PE-Cy5.5-conjugated anti-CD11c (N418), PE-Cy7-conjugated anti-CD8 (53.6.7), APC-conjugated anti-CD103 (2E7), eFluor 605NC-conjugated anti-CD45 (30-F11) and eFluor 650NC-conjugated anti-MHC class II (I-A/I-E) (M5/114.15.2) are from eBioscience. FITC-conjugated anti-CD3 (145-2C11) and anti-CD49b (DX5) are from BioLegend. The LIVE/DEAD Fixable Aqua Dead Cell Stain Kit is from Molecular Probes.

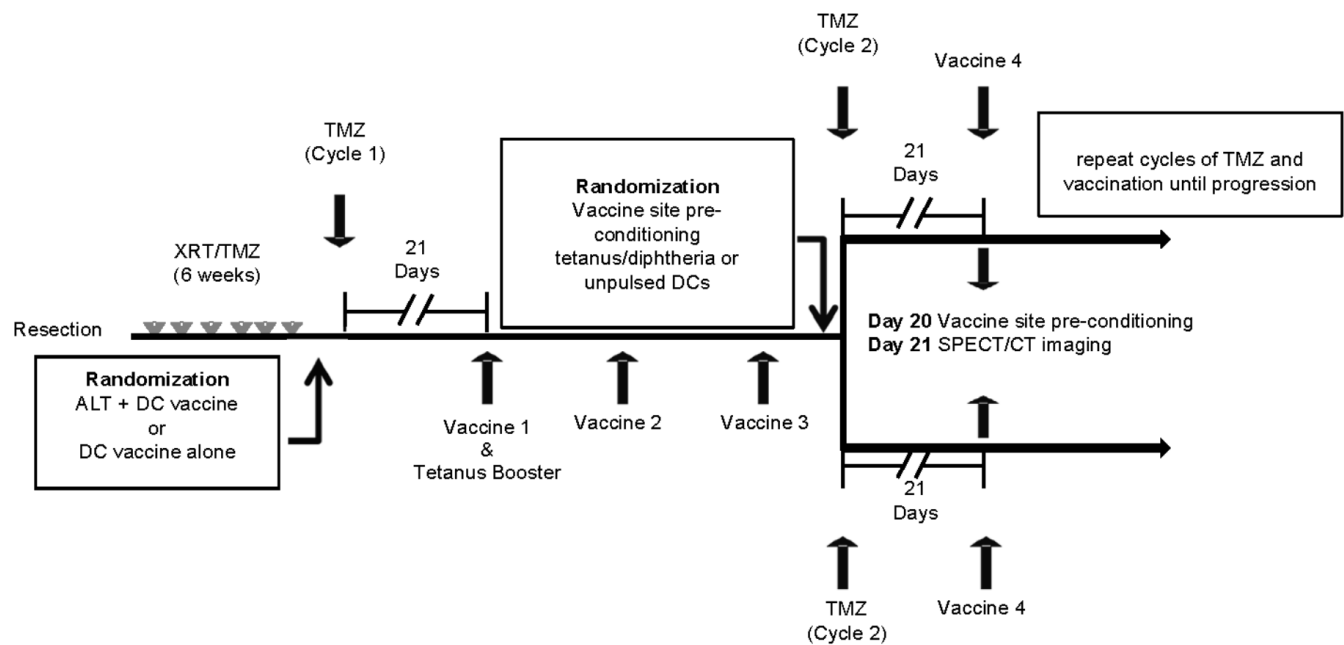
Serum cytokine and chemokine analysis. Peripheral blood was collected 24 h after vaccine site pre-conditioning before DC vaccination. For patients, blood was collected in 10 ml venous collection tubes (BD), allowed to clot, spun at 1,170g for 15 min, and serum was stored at -190 °C. For mice, blood was collected in microtainer tubes (BD) allowed to clot for 30 min, spun at 8,000g for 5 min, and serum was stored at -80 °C. Multiplex cytokine and chemokine kits were used for patient and mouse studies (cytokines and chemokines of interest for human, Affymetrix and Millipore: EPX080-10007-901, EPX010-12121-901, EPX010-12125-9, EPX010-10287-901, HCYTOMAG-60K-01 MDC; for mouse: Affymetrix and Millipore: EPX090-20821-901 ProcartaPlex 9 plex, MCYP3MAG-74K-01 MDC) following the manufacturer's instructions.

Expression of chemokines CCL3 and CCL21 in mice. Female 6–8-week-old C57BL/6 or Ccl3^{-/-} mice were immunized with Td as described above. Twenty-four

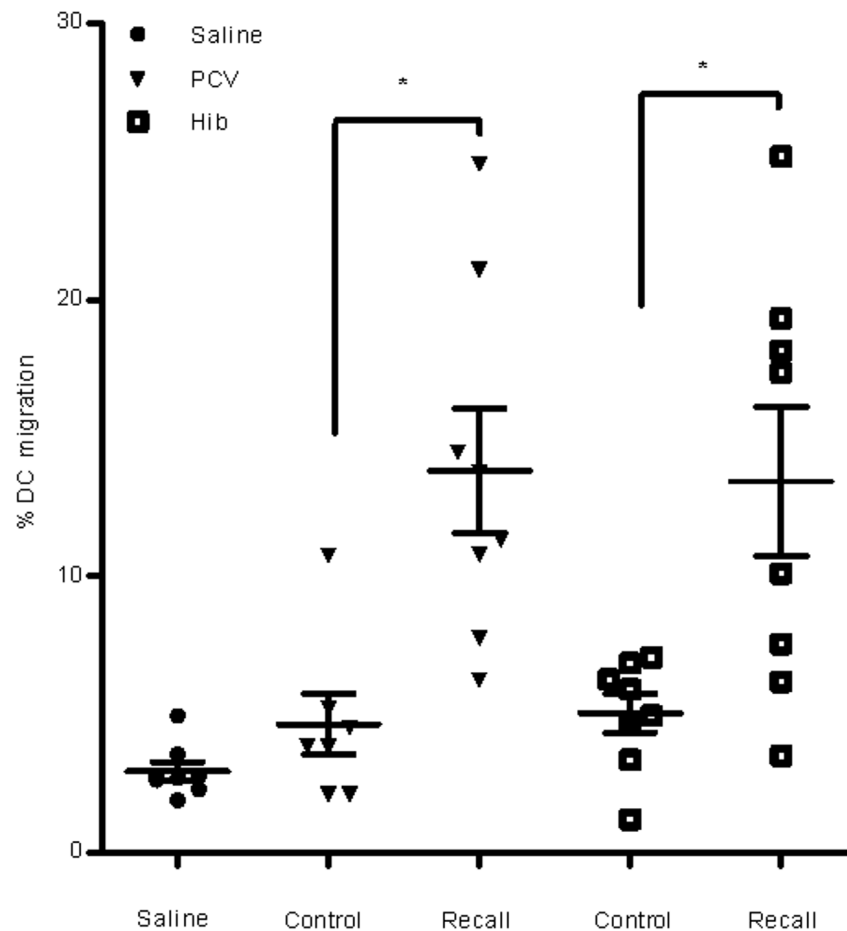
hours after Td pre-conditioning, both left and right skin sites and inguinal lymph nodes were collected. For protein isolation, skin and lymph node samples were placed in pre-loaded bead lysis Eppendorf tubes (Next Advance) containing RIPA buffer (Sigma) with protease inhibitor cocktail tablets (Mini Complete Protease Inhibitor Cocktail Tablets, Roche Applied Science). Homogenization was performed with the Bullet Blender at 4 °C. Supernatants were collected by centrifugation, and chemokines were quantified by ELISA. Quantikine kits (R&D Systems) were used for CCL3, and RayBiotech ELISA kits were used for CCL21. Corresponding samples were run for total protein concentration using the Bradford assay. CCL3 and CCL21 concentrations were normalized across samples and expressed as picograms or nanograms per milligram of total skin or lymph node protein.

Statistical analysis. Statistics were reviewed by biostatisticians and tested as described in figure legends. Cox proportional hazard models were used to evaluate DC migration and clinical outcomes. The log-rank test was used to compare Kaplan–Meier survival curves with censored patient data. For *in vivo* DC migration and VDLN studies, individual patient and mouse inguinal lymph nodes were treated as separate biological replicates, based on the underlying assumption that unilateral vaccine site pre-conditioning may preferentially affect local draining lymph nodes over non-draining sites. An unpaired two-sample Student's *t*-test was used for two-group comparisons. Paired *t*-tests were used for comparisons between lymph nodes in the same host. One-way ANOVA was used to assess differences among three or more groups with post-hoc Tukey *t*-tests for two-group comparisons. Wilcoxon rank sum analyses were conducted for pairwise comparisons in serum cytokine/chemokine panels. Signed rank tests were used to evaluate fold increase in chemokine levels. For tumour growth curves, a mixed effects linear model was employed using log-transformed curves and *F*-test for pairwise comparisons of regression line slopes and mean tumour volumes on the first day of detectable tumours (*y*-intercept). Repeated measures for calculation of slopes incorporated time between detectable tumour until considerable dropout occurred (maximal tumour size, ulceration, or death). Mean tumour volumes at final time points when the entire control cohort expired were compared between two groups using an unpaired two-sample student's *t* test. Asterisks indicate level of significance (**P* < 0.05, ***P* ≤ 0.01, ****P* < 0.001, *P* > 0.05 not significant). No statistical methods were used to predetermine sample size.

- McLendon, R. E. *et al.* Immunohistochemical detection of the DNA repair enzyme O6-methylguanine-DNA methyltransferase in formalin-fixed, paraffin-embedded astrocytomas. *Lab. Invest.* **78**, 643–644 (1998).
- Nair, S., Archer, G. E. & Tedder, T. F. Isolation and generation of human dendritic cells. *99*, 7.32:7.32.1–7.32.23 *Curr. Protoc. Immunol.* (2012).
- Thurner, B. *et al.* Generation of large numbers of fully mature and stable dendritic cells from leukapheresis products for clinical application. *J. Immunol. Methods* **223**, 1–15 (1999).
- Therasse, P. *et al.* New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J. Natl Cancer Inst.* **92**, 205–216 (2000).
- Inaba, K., Swiggard, W. J., Steinman, R. M., Romani, N. & Schuler, G. Isolation of dendritic cells. *Curr. Protoc. Immunol.* **25**, 3.7:3.7.1–3.7.15 (2001).
- Nakano, H. *et al.* Blood-derived inflammatory dendritic cells in lymph nodes stimulate acute T helper type 1 immune responses. *Nature Immunol.* **10**, 394–402 (2009).
- Sanchez-Perez, L. *et al.* Potent selection of antigen loss variants of B16 melanoma following inflammatory killing of melanocytes *in vivo*. *Cancer Res.* **65**, 2009–2017 (2005).
- Daniels, G. A. *et al.* A simple method to cure established tumors by inflammatory killing of normal cells. *Nature Biotechnol.* **22**, 1125–1132 (2004).
- Fidler, I. J. Biological behavior of malignant melanoma cells correlated to their survival *in vivo*. *Cancer Res.* **35**, 218–224 (1975).

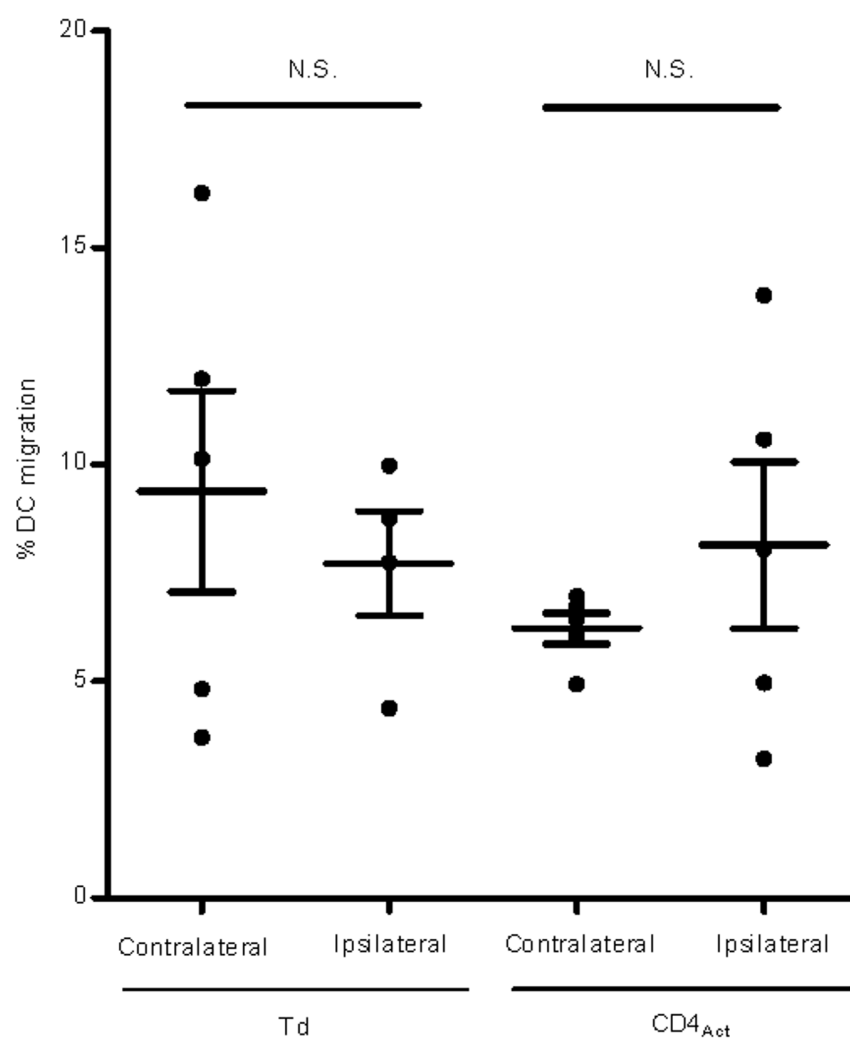


Extended Data Figure 1 | Schema of clinical trial. SPECT/CT, single photon emission computed tomography/computed tomography; TMZ, temozolomide; XRT, external beam radiotherapy.



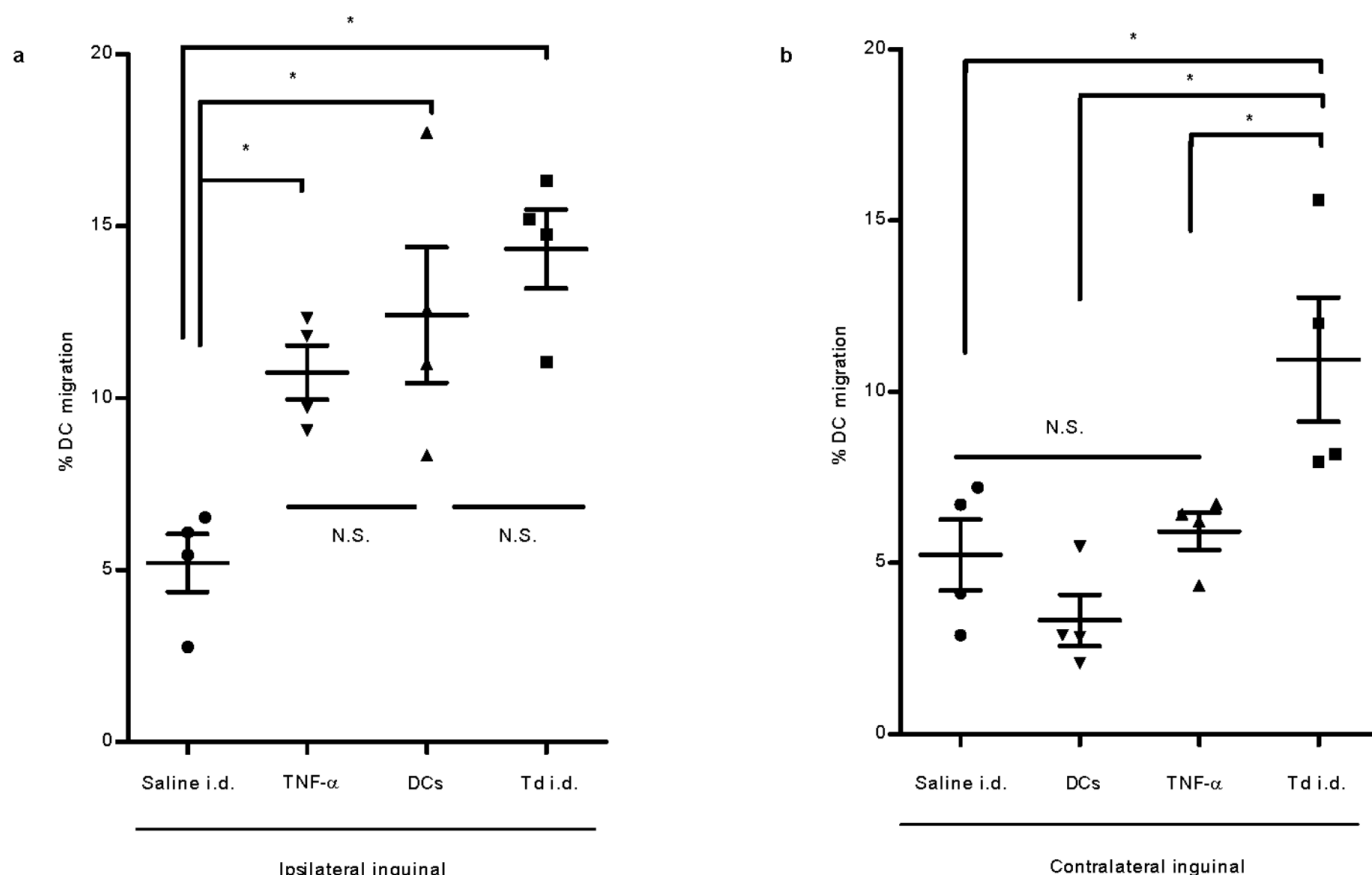
Extended Data Figure 2 | Recall responses induced by other CD4⁺ T-cell-dependent protein antigens increase DC migration to VDLNs. Primary immunization and vaccine site pre-conditioning with CD4⁺ T-cell-dependent protein antigens increase DC migration to VDLNs. Mice were immunized with either Haemophilus b conjugate (Hib) or pneumococcal 13-valent conjugate (PCV) intramuscularly, and 2 weeks later received vaccine site pre-conditioning with the recall antigen (recall) or saline (control). A separate

cohort of mice received saline only throughout the immunization schedule (saline). Scatter plot shows biological replicates of individually processed right and left iLN per mouse (4 mice per group). Percentage migration of RFP⁺ DCs to VDLNs; one-way ANOVA, $P < 0.0001$; post-hoc Tukey t -test, PCV control versus recall, $P < 0.05$, Hib control versus recall, $P < 0.05$. Representative of $n = 3$ experiments; mean \pm s.e.m.



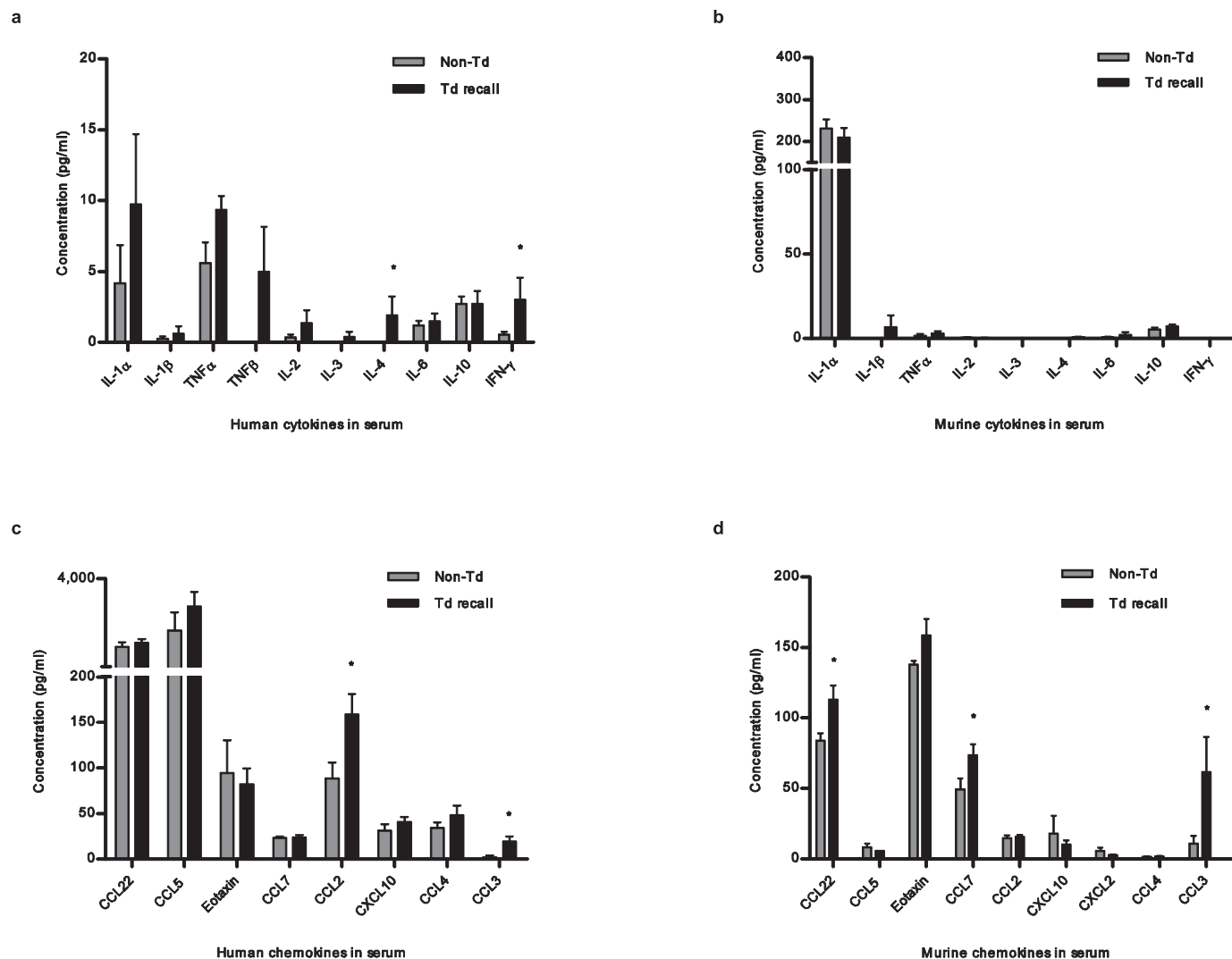
Extended Data Figure 3 | Bilateral migration of OVA-DCs after Td pre-conditioning or Td-activated CD4⁺ T-cell transfer. Uptake of injected DCs to right and left iLNs 48 h after DC vaccination in Td-immune mice receiving Td pre-conditioning or naive mice administered Td-activated CD4⁺ T cells.

Scatter plot shows biological replicates of individually processed right and left iLN per mouse (5 mice per group). CD4_{act} ipsilateral versus contralateral, paired *t*-test, *P* = 0.41. Representative of *n* = 4 experiments; mean ± s.e.m.



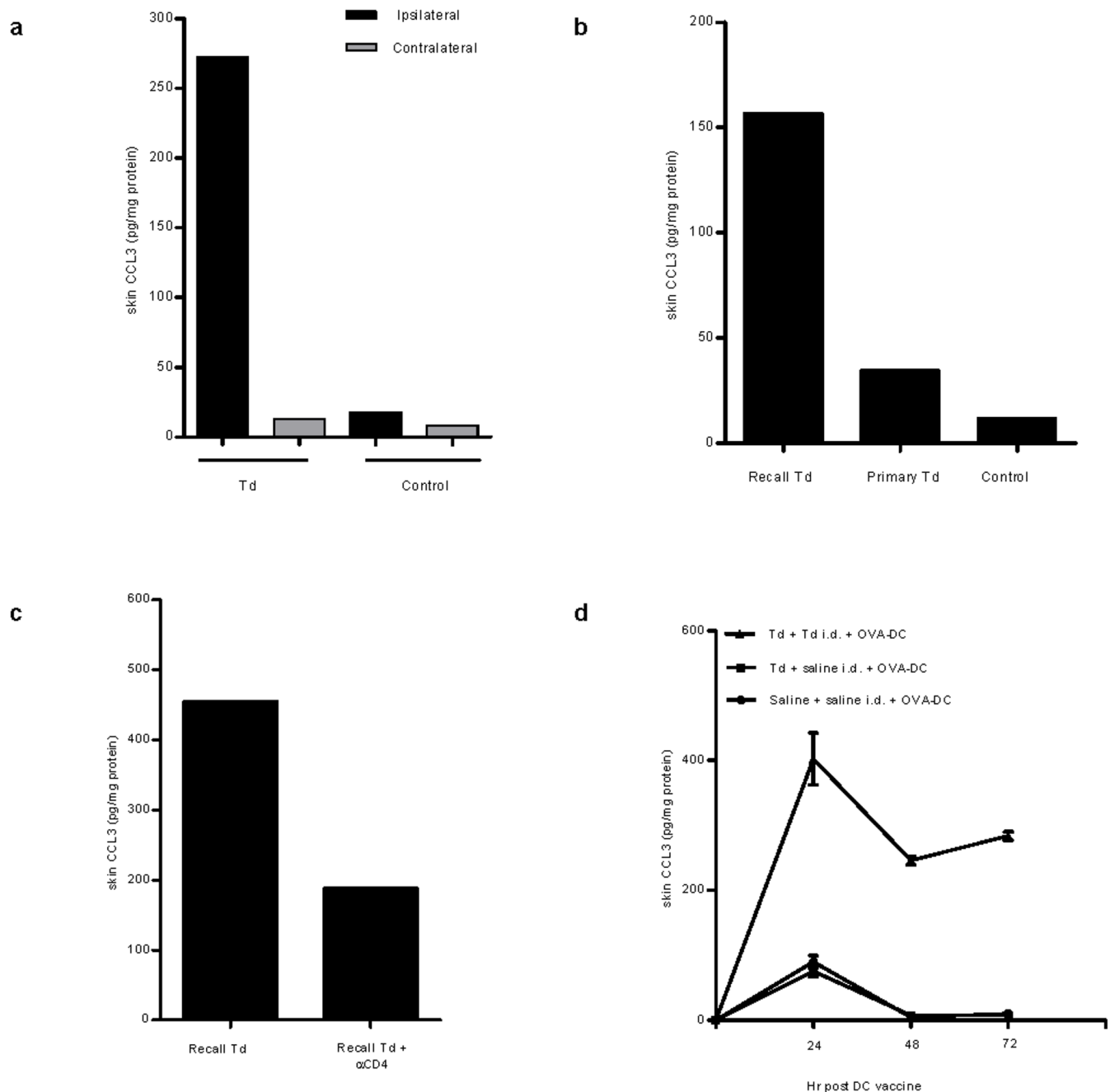
Extended Data Figure 4 | Unilateral pre-conditioning with unpulsed DCs or TNF- α results in increased DC homing to ipsilateral draining inguinal lymph nodes. Td-immune mice pre-conditioned with Td or saline before administration of OVA RNA-pulsed DC vaccine. Separate cohorts of naive mice received either 1×10^6 unpulsed DCs or 30 ng TNF- α on one side of the groin 24 h before the bilateral RFP⁺ DC vaccine. DC migration was quantified 24 h after vaccination. **a**, DC migration to ipsilateral lymph nodes (one-way ANOVA, $P = 0.0018$; post-hoc Tukey t -test, saline i.d. versus Td i.d., $P = 0.007$;

saline i.d. versus TNF- α , $P < 0.05$; saline i.d. versus DCs, $P < 0.05$; Td i.d. versus TNF- α , $P = 0.042$; DCs versus Td i.d. and DCs versus TNF- α , $P > 0.05$. **b**, DC migration to contralateral lymph nodes; one-way ANOVA, $P = 0.003$; post-hoc Tukey t -test, saline i.d. versus DCs or TNF- α , $P > 0.05$; Td i.d. versus TNF- α , DCs or saline i.d., $P < 0.05$. n values are biological replicates of individually processed right and left iLNs per mouse (4 mice per group). Representative of $n = 3$ experiments; mean \pm s.e.m.



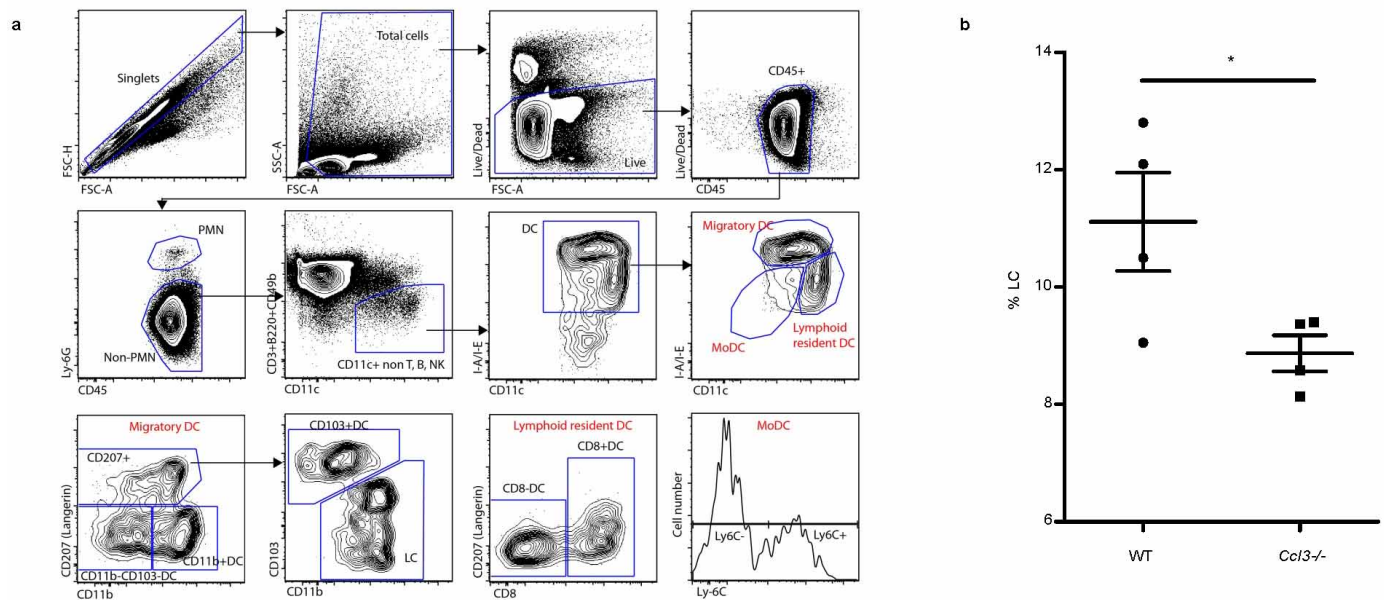
Extended Data Figure 5 | Serum cytokine and chemokine profile after Td pre-conditioning in patients and mice. **a**, Serum cytokine panel of patients following vaccine site pre-conditioning with Td or unpulsed DCs; Wilcoxon rank sum, IFN- γ and IL-4, $P < 0.05$ ($n = 6$ patients). **b**, Similar to **a** in mice; Wilcoxon rank sum, all comparisons, $P > 0.05$ (Td recall $n = 5$, non-Td $n = 6$). **c**, Patient serum chemokines after vaccine site pre-conditioning. Patient

CCL2 and CCL3 in Td recall (Td, $n = 6$) versus non-Td (unpulsed DC, $n = 5$); one-way ANOVA and Wilcoxon rank sum, $P < 0.05$. **d**, Mouse CCL22, CCL7 and CCL3 in Td recall (Td, $n = 8$ mice) and non-Td (saline, $n = 8$ mice); one-way ANOVA and Wilcoxon rank sum, $P < 0.05$. For **a–d**, individual values represent biological replicates; mean \pm s.e.m.



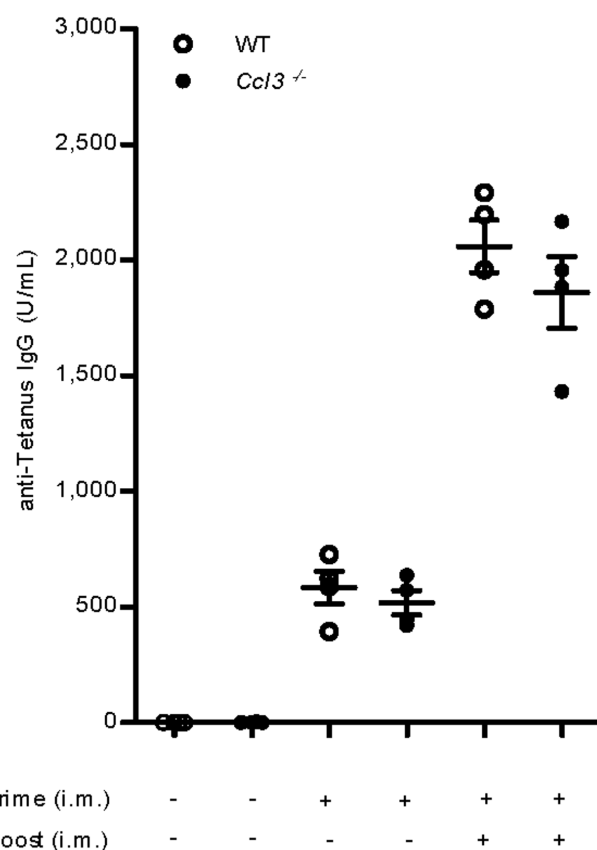
Extended Data Figure 6 | Td vaccine site pre-conditioning results in CCL3 upregulation in Td-immune hosts. **a**, CCL3 production in skin site after Td pre-conditioning (Td ipsilateral versus contralateral). Representative of four independent experiments. **b**, CCL3 production in skin after Td recall response. **c**, CCL3 induction at skin site is abrogated with previous host depletion of CD4⁺ T cells. Bars in **a–c** represent CCL3 protein detected in skin sites from

$n = 2$ mice with $n = 2$ technical replicates performed per mouse. **d**, CCL3 remains increased at the Td pre-conditioning site in the skin after DC vaccination (24, 48 and 72 h, one-way ANOVA, $P = 0.0001$, Td plus Td i.d. and OVA-DC versus Td plus saline i.d. and OVA-DC, and saline plus saline i.d. and OVA-DC, $P < 0.05$, post-hoc Tukey t -test). Individual values represent biological replicates from $n = 4$ mice; mean \pm s.e.m.

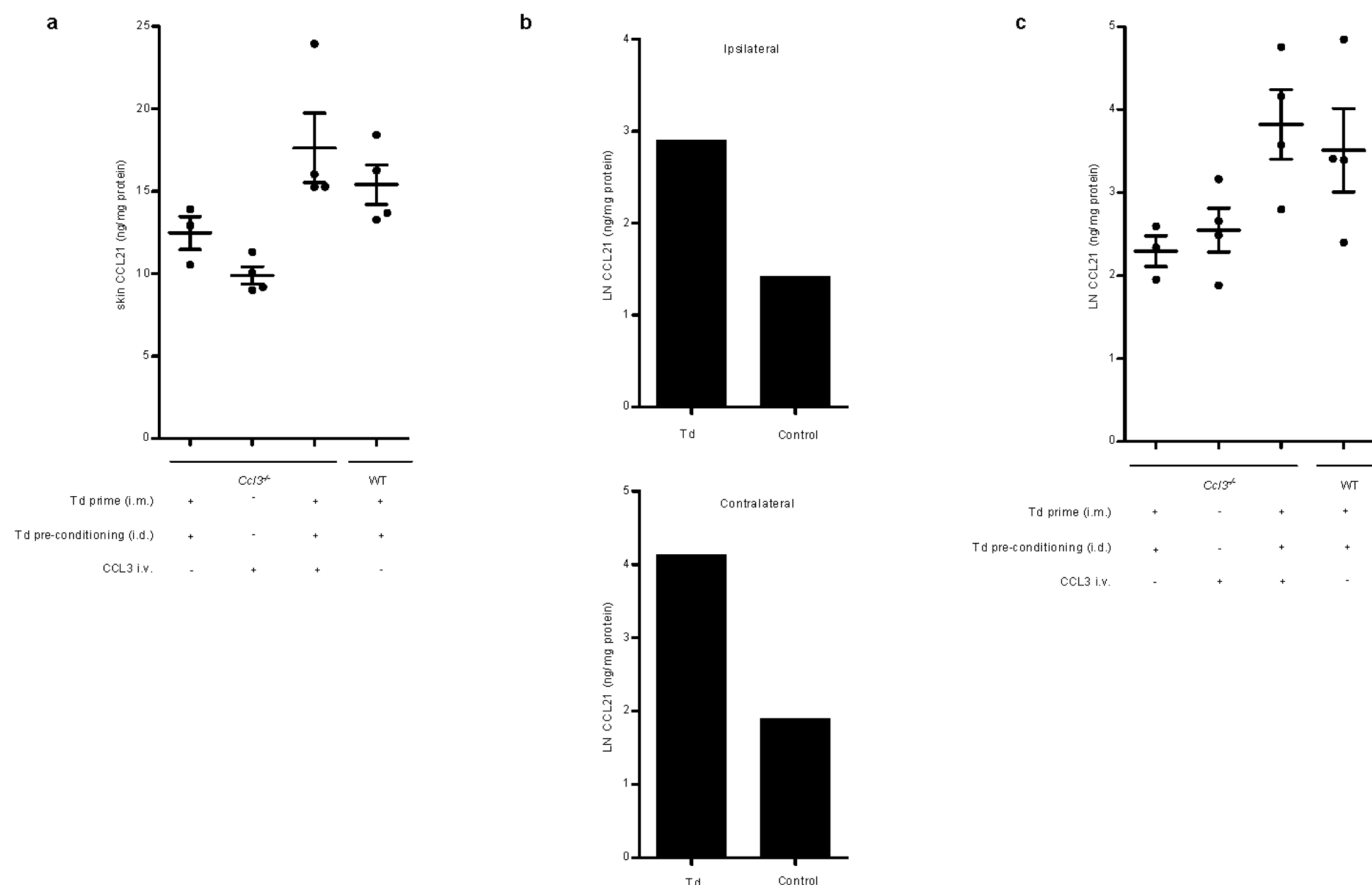


Extended Data Figure 7 | Migratory DC subsets in wild-type and *Ccl3*^{-/-} mice after induction of Td recall responses. Both wild-type and *Ccl3*^{-/-} mice were first immunized with Td and then challenged with Td pre-conditioning. Migration of endogenous DC subsets to inguinal lymph nodes contralateral to the site of Td pre-conditioning was assessed at days 4 and 8 after Td administration. **a**, Gating strategy used to quantify DC subsets in

inguinal lymph nodes after skin pre-conditioning with Td. LC, Langerhans cells; MoDC, monocyte-derived DCs. **b**, Day-8 migration of LC population to non-draining inguinal lymph nodes in *Ccl3*^{-/-} hosts is reduced in absence of CCL3; two-sample *t*-test, *P* = 0.046. Representative of three experiments. Individual values represent biological replicates from *n* = 4 mice; mean ± s.e.m.



Extended Data Figure 8 | Anti-tetanus toxoid memory responses are induced and maintained in wild-type and *Ccl3*^{-/-} mice throughout Td priming and boosting. Wild-type and *Ccl3*^{-/-} mice primed and boosted with Td. Serum from immunized mice was collected 2 weeks after each immunization before the next booster vaccine (for each boosting phase, wild-type versus *Ccl3*^{-/-}, two-sample *t*-test, *P* > 0.05). i.m., intramuscular. Scatter plot showing averaged values from *n* = 4 mice with *n* = 2 technical replicates performed per mouse. Representative of three experiments; mean ± s.e.m.



Extended Data Figure 9 | CCL21 levels in Td pre-conditioning skin sites and draining lymph nodes of wild-type and *Ccl3*^{-/-} mice. **a**, CCL21 levels in skin site of *Ccl3*^{-/-} hosts after induction of Td recall response and CCL3 administration. Mixed model accounting for within-mouse correlation of measurements, *F*-test, $P < 0.001$; pairwise comparisons, WT Td plus Td i.d. ($n = 4$) versus *Ccl3*^{-/-} Td plus Td i.d. ($n = 3$), $P = 0.049$; *Ccl3*^{-/-} Td plus Td i.d. and CCL3 i.v. ($n = 4$) versus *Ccl3*^{-/-} Td plus Td i.d. and versus *Ccl3*^{-/-} plus CCL3 i.v. ($n = 4$), $P = 0.044$ and $P = 0.0045$, respectively. Scatter plot shows averaged values with $n = 2$ technical replicates performed per mouse. **b**, Bilateral iLN CCL21 levels in wild-type mice after Td recall with skin site pre-conditioning. Bars represent CCL21 protein within iLN ipsilateral and contralateral to the side of Td pre-conditioning from $n = 2$ mice with $n = 2$

technical replicates performed per iLN. Representative of three experiments. **c**, Increased lymph node CCL21 in *Ccl3*^{-/-} hosts after CCL3 reconstitution and induction of Td recall response (all two group comparisons). CCL21 levels in bilateral iLNs of *Ccl3*^{-/-} hosts after induction of Td recall response and CCL3 administration. Mixed model accounting for within-mouse correlation of measurements, *F*-test, $P < 0.001$; pairwise comparisons, WT Td plus Td i.d. ($n = 4$ iLNs) versus *Ccl3*^{-/-} Td plus Td i.d. ($n = 3$ iLNs), $P = 0.045$; *Ccl3*^{-/-} Td plus Td i.d. and CCL3 i.v. ($n = 4$ iLNs) versus *Ccl3*^{-/-} Td plus Td i.d. and versus *Ccl3*^{-/-} plus CCL3 i.v. ($n = 4$ iLNs), $P = 0.0066$ and $P = 0.026$, respectively. Scatter plot shows averaged values with $n = 2$ technical replicates performed per lymph node sampled. In **a** and **c**, mean \pm s.e.m.

Extended Data Table 1 | Clinical trial patient characteristics

Patient	Sex	Age	Biopsy location	Race	KPS	WHO score	MMSE	IDH1	MGMT	MGMT promoter methylation	Vaccine site pre-conditioning	PFS (randomization)	OS (randomization)	PFS (diagnosis)	OS (diagnosis)	EORTC predicted median OS	EORTC O-E	RPA class	RPA O-E
1	F	46	R. frontal	W	100	0	29	-	N/A	N/A	Td	8.8	19.2	15.4	25.7	26.8*	-1.1	III	7.8
2	F	32	R. frontal	W	100	0	29	+	-	-	DC	21.7	35.2	27.8	41.3	13.0††	28.3††	III	23.4
3	M	62	R. temporal	W	90	1	29	-	-	-	DC	4.9	10.9	11.2	17.3	9.8	7.5	IV	6.2
4	M	71	L. frontoparietal	W	90	1	30	-	-	+	Td	41.4	41.4	47.3†	47.3†	16.4	30.9	IV	36.2
5	M	43	L. temporal	W	90	1	28	-	-	+	DC	2.5	12.2	10.0	19.7	16.4	3.3	III	1.8
6	F	59	R. temporal	W	80	1	30	-	-	-	Td	39.5	39.5	45.0†	45.0†	9.8	35.2	IV	33.9
7	M	75	R. temporal	B	90	1	26	-	+	+	Td	9.5	14.7	15.4	20.6	11.8	8.8	IV	9.5
8	F	58	L. parietal	H	90	1	22	-	+	N/A	DC	2.5	13.4	9.1	20.0	13.1*	6.9	IV	8.9
9	F	30	R. temporal	W	90	1	29	N/A	+	N/A	Td	36.6	36.6	44.1†	44.1†	21.3*	22.8	III	26.2
10	F	66	L. frontoparietal	W	80	1	24	-	+	-	DC	3.9	9.5	10.4	16.0	6.9	9.1	IV	4.9
11	F	28	L. frontotemporal	W	100	0	28	-	-	-	DC	6.1	7.4	12.5	13.8	13.0	0.8	III	-4.1
12	M	59	R. posteroparietal	W	90	1	29	§	§	§	DC	§	§	5.1	19.0	18.1*	0.9	IV	7.9
13	M	71	L. posterotemporal	W	90	1	21	-	+	+	Td	12.7	15.2	18.5	20.9	11.8	9.1	IV	9.8
Median DC		58.5			90	1	28					4.4	11.6	10.8	18.5	13.0††	6.9††	IV	6.2
Median Td		65			90	1	29					NE	NE	NE	NE	14.1	>16	IV	>18

Demographic and prognostic factors of newly diagnosed GBM patients with vaccine site pre-conditioning randomization strategy and corresponding progression-free and overall survival from the time of surgery and from randomization to pre-conditioning. Observed and predicted survival times are expressed in months. Median values are shown for Td and unpulsed DC cohorts. Predicted median overall survival for RPA class yielded 17.9 months for class III and 11.1 months for class IV based on recursive partition analysis¹⁵. Model 3 of the EORTC scoring system that incorporated *MGMT* promoter methylation status, MMSE score, and WHO performance status was used to generate predicted median survival rates.

* Model 2 of EORTC scoring system if methylation status unavailable.

† No progression.

‡ Alive.

§ Patient progressed before time of randomization for vaccine site pre-conditioning.

EORTC, European Organization for Research and Treatment of Cancer; L, left; IDH1, isocitrate dehydrogenase type1; KPS, Karnofsky performance status; MGMT, O⁶-methylguanine-DNA methyltransferase; MMSE, mini-mental state examination; NA, tissue not available; NE, not estimable; O-E, observed - expected survival months; OS, overall survival; PFS, progression-free survival; R, right; RPA, recursive partitioning analysis; WHO, World Health Organization.

A human tRNA synthetase is a potent PARP1-activating effector target for resveratrol

Mathew Sajish¹ & Paul Schimmel^{1,2}

Resveratrol is reported to extend lifespan^{1,2} and provide cardio-neuro-protective³, anti-diabetic⁴, and anti-cancer effects^{3,5} by initiating a stress response² that induces survival genes. Because human tyrosyl transfer-RNA (tRNA) synthetase (TyrRS) translocates to the nucleus under stress conditions⁶, we considered the possibility that the tyrosine-like phenolic ring of resveratrol might fit into the active site pocket to effect a nuclear role. Here we present a 2.1 Å co-crystal structure of resveratrol bound to the active site of TyrRS. Resveratrol nullifies the catalytic activity and redirects TyrRS to a nuclear function, stimulating NAD⁺-dependent auto-poly-ADP-ribosylation of poly(ADP-ribose) polymerase 1 (PARP1). Downstream activation of key stress signalling pathways are causally connected to TyrRS–PARP1–NAD⁺ collaboration. This collaboration is also demonstrated in the mouse, and is specifically blocked *in vivo* by a resveratrol-displacing tyrosyl adenylate analogue. In contrast to functionally diverse tRNA synthetase catalytic nulls created by alternative splicing events that ablate active sites⁷, here a non-spliced TyrRS catalytic null reveals a new PARP1- and NAD⁺-dependent dimension to the physiological mechanism of resveratrol.

Human TyrRS is a homodimer of a 528 amino-acid polypeptide that harbours an appended eukaryote-specific carboxy (C)-terminal EMAP-II domain (Fig. 1a)^{8,9}. High-resolution structures of the catalytic unit of *Hs* TyrRS, known as mini-TyrRS, and the C-domain have been determined^{10,11}. We found resveratrol (RSV) strongly inhibited TyrRS with an inhibition constant (K_i) value of 22 μM (Extended Data Fig. 1a–c). Crystallization of *Hs* mini-TyrRS with RSV and, separately, with tyrosine yielded co-crystal structures (at 2.1 Å) (Fig. 1b, Extended Data Fig. 2a, b and Extended Data Table 1; Protein Data Bank (PDB) accession numbers 4Q93 and 4QBT).

While the phenolic ring of RSV and of the tyrosine have the same disposition in the respective co-crystals, accommodation of the *cis*-conformation of the dihydroxy ring of RSV forces a local structural change near the linker to the C-domain (Fig. 1b and Extended Data Fig. 2a). An RSV-promoted conformational change in TyrRS may drive

the predominant *trans* RSV (in solution) into a *cis* conformation (Extended Data Fig. 2c, d).

Associated with a previous study¹², a distinct TyrRS–PARP1 interaction was observed. PARP1 is a major modulator of NAD⁺ metabolism and its related signalling¹³. Because RSV acts through NAD⁺-dependent proteins¹⁴, the TyrRS–PARP1 interaction was further studied. Given that RSV treatment elicits a stress response², serum starvation was used to mimic a general ‘stand-alone’ stress condition so that common signalling pathways, if any, between RSV treatment and a general stress condition, could be compared *ex vivo*. Either serum starvation or RSV treatment promoted nuclear translocation of endogenous TyrRS in HeLa cells (Fig. 1c). Translocation was observed under different stress conditions (heat shock and endoplasmic reticulum stress; Extended Data Fig. 3a), suggesting that TyrRS is a general stress transducer. Nuclear translocation of endogenous TyrRS was concomitant with strong auto-PARylation of PARP1 (PARP1^{PAR}) (Fig. 1c and Extended Data Fig. 3a).

Ex vivo RSV also strongly promoted association of TyrRS with PARP1, and robust auto-poly-ADP-ribosylation of PARP1 (Extended Data Fig. 3b). Effects of RSV were blocked by a Tyr-AMP analogue (Tyr-SA (5'-O-[N-(9L-tyrosyl) sulphamoyl] adenosine)), but not by Gly-SA (a control targeting GlyRS) (Extended Data Fig. 3b, c). Similar, but less pronounced, PARylation was seen with serum starvation. Enhanced PARylation correlated with increased amounts of TyrRS in the nucleus, which occurred upon serum starvation. Thus, both serum starvation and RSV promoted nuclear translocation of TyrRS and activation of PARP1.

Cell lysates treated with the PARG hydrolyase and its hydrolase-inactive mutant supported the idea that TyrRS preferentially bound to non-PARylated PARP1 (Extended Data Fig. 3d, e). TyrRS interacted specifically with the C-domain of PARP1 (CT-PARP1) (Extended Data Fig. 3f). Neither mini-TyrRS nor the TyrRS C-domain interacted with PARP1; only full-length native TyrRS bound PARP1 (Extended Data Fig. 3g, h). In the absence of RSV, concentration-dependent activation of PARP1 by TyrRS was observed *in vitro* (Fig. 2a, top, and Extended

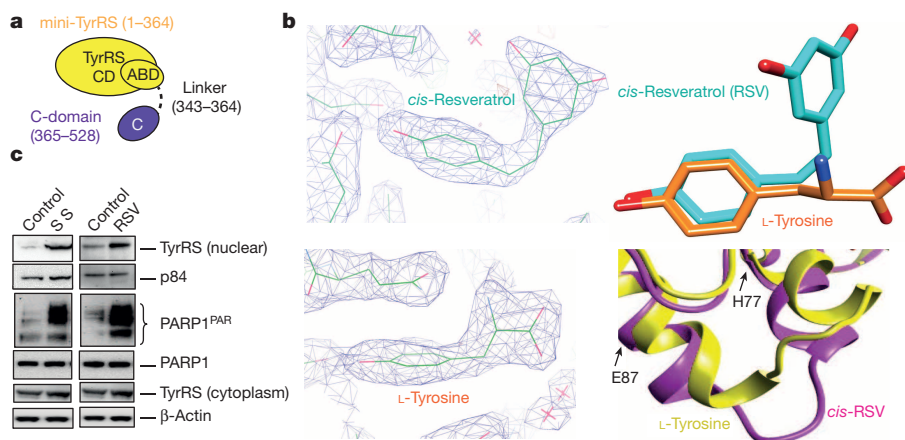


Figure 1 | Resveratrol binds at the active site of TyrRS. **a**, Cartoon illustration of the domain organization of human TyrRS. Both domains are connected by a linker of approximately 20 amino acids. **b**, Left, Electron density of co-crystal X-ray structures (2.1 Å) of TyrRS bound to *cis*-resveratrol and to L-tyrosine. Right, resveratrol induced a local conformational change relative to bound tyrosine at the active site. **c**, Both serum starvation and resveratrol treatment (5 μM) facilitated the nuclear translocation of TyrRS with a concomitant increase in the PARylation of PARP1.

¹The Skaggs Institute for Chemical Biology, The Scripps Laboratories for tRNA Synthetase Research, Department of Molecular and Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA. ²The Scripps Florida Research Institute, 130 Scripps Way, Jupiter, Florida 33458, USA.

Data Fig. 4a, b). RSV enhanced *in vitro* auto-PARylation with the half-maximal effect at roughly 10 nM (Fig. 2a, middle), well below the K_i (about 22 μ M) in Extended Data Fig. 1a–c. Thus, PARP1 may alter the apparent affinity of RSV for TyrRS. Also, concentration-dependent quenching of PARylation of PARP1 by Tyr-SA was evident (Fig. 2a, bottom). Lastly, while broken DNA normally activates PARP1 (ref. 13), Tyr-SA did not interfere with this DNA-dependent-pathway of PARP1 activation *in vitro* (Extended Data Fig. 4c). Therefore, TyrRS–RSV activation of PARP1 is distinct.

Ectopically expressed TyrRS in HeLa cells for 0–24 h caused progressive increase in cellular concentrations of the synthetase (Fig. 2b, top) and a correlated progressive increase in PARP1^{PAR} (Fig. 2b, top). A TyrRS mutant (TyrRS-dNLS), with reduced nuclear localization⁶, reduced activation of PARP1 (Extended Data Fig. 4d). Effects of RSV *ex vivo* at various concentrations (Fig. 2b, middle) and of Tyr-SA (Fig. 2b, bottom) mirrored those seen *in vitro* (Fig. 2a, middle and bottom).

RSV and L-tyrosine produce different bound conformations of TyrRS, seen locally in the high-resolution structures of the two co-crystals (Fig. 1b and Extended Data Fig. 2a). These differences could affect the disposition of the C-domain needed for the interaction of TyrRS with PARP1. This domain is tethered to the mini-TyrRS catalytic unit by a flexible linker (Fig. 2c, left). By breaking the Y341 OH–H bond to a backbone carbonyl oxygen, a Y341A mutation releases tight tethering

of the C-domain to the catalytic domain¹⁵ and its reorientation¹⁶ (Fig. 2c, right).

In the absence of RSV, Y341A-TyrRS showed robust interaction with and activation of PARP1 *in vitro* (Fig. 2d). If Y341A is shifted more towards a conformation needed to bind to PARP1, then it would be more sensitive than wild-type TyrRS to low concentrations of RSV. This expectation was fulfilled (Extended Data Fig. 4e). TyrRS–Tyr may thus have a conformation that prevents interaction with PARP1, while TyrRS–RSV has a distinct conformation that binds PARP1. In the absence of either ligand, a dynamic equilibrium populates both forms (Fig. 2c).

A working model of downstream signalling markers was developed as detailed in Methods (Fig. 3a). Effects from serum starvation and RSV administration involve cascades through protein acetylation and phosphorylation effects of serum starvation and of RSV, which have not been linked to TyrRS nor activation of PARP1. Under conditions of serum starvation or RSV administration, the response in HeLa cells of PARylated and acetylated proteins, together with the aforementioned cell-signalling proteins, was monitored for up to 1 h. Either RSV or serum-starvation promoted production of AcK-Tip60 and AcK-ATM (Fig. 3b). In a temporal response from 0 to 8 h, either or both of serum starvation or 1 μ M RSV promoted increased levels of whole-cell PARylated

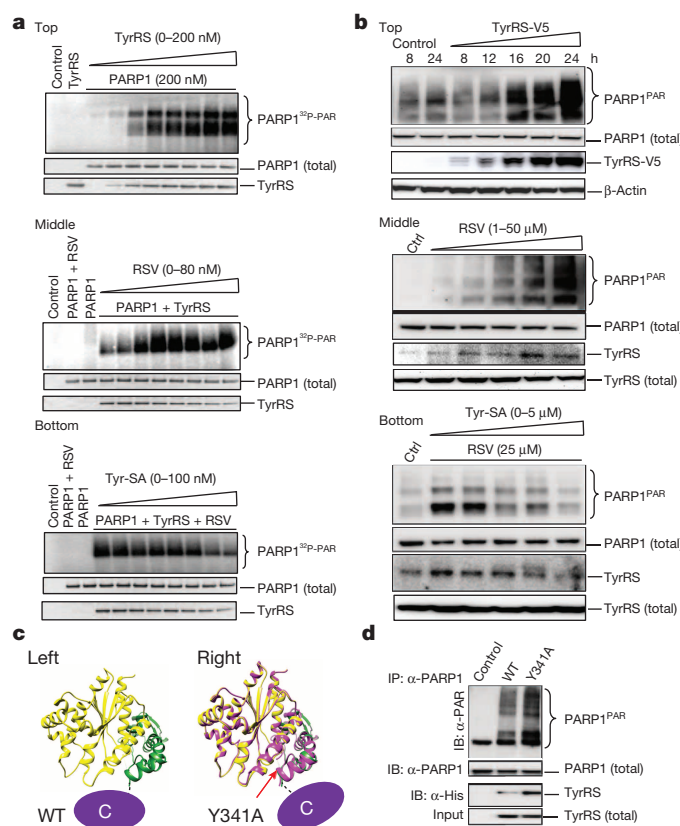


Figure 2 | TyrRS facilitates the activation of PARP1 in an active-site-dependent manner. **a**, Top: TyrRS activates PARP1 in an *in vitro* assay. Middle: resveratrol potentiates TyrRS-mediated activation of PARP1. Bottom: Tyr-SA blocks the resveratrol-mediated activation of PARP1. **b**, Top: TyrRS-V5 overexpression activates PARP1 in HeLa cells in a concentration-dependent manner. Middle: resveratrol treatment activates PARP1 in HeLa cells and enhances TyrRS interaction with PARP1. Ctrl, control. Bottom: Tyr-SA blocks the resveratrol-mediated interaction of TyrRS and activation of PARP1. **c**, Cartoon illustration of the C-domain disposition in TyrRS (left) and Y341ATyrRS (right). WT, wild type. **d**, Y341ATyrRS enhances its interaction and activates PARP1 compared with wild type. IP, immunoprecipitation; IB, immunoblot.

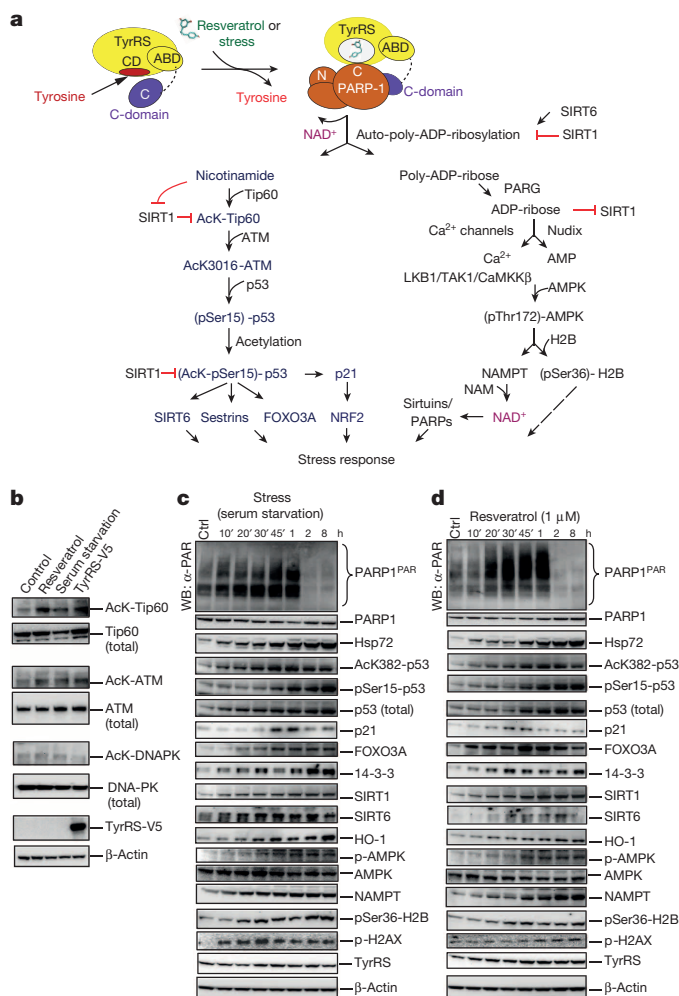


Figure 3 | Resveratrol and serum starvation mimic similar downstream signalling events mediated by PARP1 activation. **a**, Cartoon illustrating the molecular basis and the integration of different signalling pathways to mediate a TyrRS/PARP1-activated stress response evoked by either resveratrol or different stress conditions. **b**, PARP1-activating conditions enhance Tip60-mediated activation of ATM. **c**, Time course (minutes and hours) of poly-ADP-ribosylation status and associated signalling events as depicted in Fig. 3a after serum starvation and **d** resveratrol (1 μ M) treatment.

proteins, Hsp72, AcK382-p53, p-AMPK, pSer36-H2B, p21, FOXO3A, 14-3-3, HO-1, NAMPT, SIRT1, and SIRT6 (Fig. 3c, d). Also, RSV activated acetylation of Tip60 *ex vivo* in a concentration-dependent manner (Extended Data Fig. 5a). As expected from previous reports that acetylation of p53 is indispensable for its activation¹⁷, a transient increase in acetylation of p53 (AcK382)¹ was evident (Fig. 3c, d), possibly through a transient inhibition of SIRT1 by the nicotinamide¹⁸ being produced. Consistently, at 15 min, an RSV concentration-dependent reduction in total NAD⁺ with concomitant production of nicotinamide and ADP-ribose was observed (Extended Data Fig. 5b, c). Thus, serum starvation and RSV rapidly increased all monitored proteins in a way that is dependent on having active PARylation. This rapid increase is consistent with work showing rapid upregulation of PARP1 target genes^{19,20}, possibly because of a broader effect of PARP1 on transcription and associated stress signalling^{13,21}.

At longer times, and continuing at low concentrations ($\leq 10 \mu\text{M}$) of RSV, similarity between responses to RSV and serum starvation was ostensibly seen. By 1 h, low concentrations of RSV elevated the NAD⁺ levels similar to those of serum-starvation conditions (Extended Data Fig. 5d). Responses of key stress-signalling markers in a more extended period of 8–24 h at 5 μM RSV were mostly similar to those of serum starvation (Extended Data Fig. 5e, f). At 5 μM RSV, levels of PARP1^{PAR} were sustained longer compared with what was observed at 1 μM RSV (compare 2 h and 8 h time points in Fig. 3d with Extended Data Fig. 5f) and we thus continued further studies at 5 μM RSV.

A short interfering RNA (siRNA) directed against PARP1 (siRNA^{PARP1}) effectively abrogated the RSV-stimulated expression of Hsp72, p-AMPK, SIRT1, FOXO3A, SESN2, NAMPT, PUMA, and SIRT6 (Fig. 4a). Also, 5 μM RSV promoted induction of BRCA1 and p14ARF (whose genes are directly regulated by PARP1 (refs 22, 23)). Consistently, the RSV-stimulated inductions of BRCA1 and p14ARF were prevented by siRNA^{PARP1} (Fig. 4a), further showing involvement of PARP1 in RSV-mediated induction of these proteins.

In Fig. 4a, RSV treatment led to enhanced activation of p53. However, in the absence of RSV, knockdown of PARP1 resulted in a background increase in the products of these p53-regulated genes. Possibly p53 was activated by the knockdown of PARP1, because removal of PARP1 would enhance background levels of DNA damage. Indeed, activation of p53, as seen by enhanced acetylation and phosphorylation of p53, was observed (Fig. 4a).

siRNA^{TyrRS}-knockdown at 5 μM RSV eliminated induction of, and dramatically reduced amounts of, PARylated PARP1, of whole acetylated proteins, and of AcK382-p53, AcK16-H4 (Tip60 activation), p-AMPK, and pSer36-H2B (AMPK activation) (Fig. 4b). (Separately, RSV (5 μM) addition did not affect the viability of HeLa cells expressing either siRNA^{TyrRS} or siRNA^{PARP1} (Extended Data Fig. 6).) In contrast, siRNA^{SIRT1}-knockdown did not affect RSV (5 μM)-mediated production of whole acetylated proteins or, among other downstream markers, activation of PARP1 and induction of SIRT6, FOXO3A, NAMPT, AcK16-H4, p-AMPK²⁴, and pSer36-H2B (Extended Data Fig. 7). Collectively, Fig. 4a, b shows that *ex vivo* TyrRS and PARP1 collaborate to activate RSV- and serum-starvation responses.

Although the RSV–TyrRS–PARP1 axis initially consumes NAD⁺, at low RSV (5 μM), NAD⁺ levels are transiently raised after 1 h owing to NAMPT activation (Extended Data Fig. 5d). At 1 h at low RSV, induction of NAD⁺ levels was abolished with either siRNA^{TyrRS} or siRNA^{PARP1} (Extended Data Fig. 8). NAD⁺ depletion by NAMPT inhibition abolished RSV-mediated induction of BRCA1, FOXO3A, NAMPT, SESN2 and SIRT6 (Fig. 4c). These results are consistent with those using inhibitors against PLC- γ and the ryanodine receptor²⁵, which inhibit PARP1 activation²⁶ and ADP-ribose-mediated calcium signalling^{25,27}, and which prevented RSV-mediated AMPK activation. Similarly, a PARP1 inhibitor prevented LKB1 activation by RSV²⁸.

Mice were injected intravenously the through tail vein with 100 μl of 10 μM RSV and, after 30 min, were euthanized. Compared with the PBS intravenous-injected control, PARylated and acetylated proteins, along

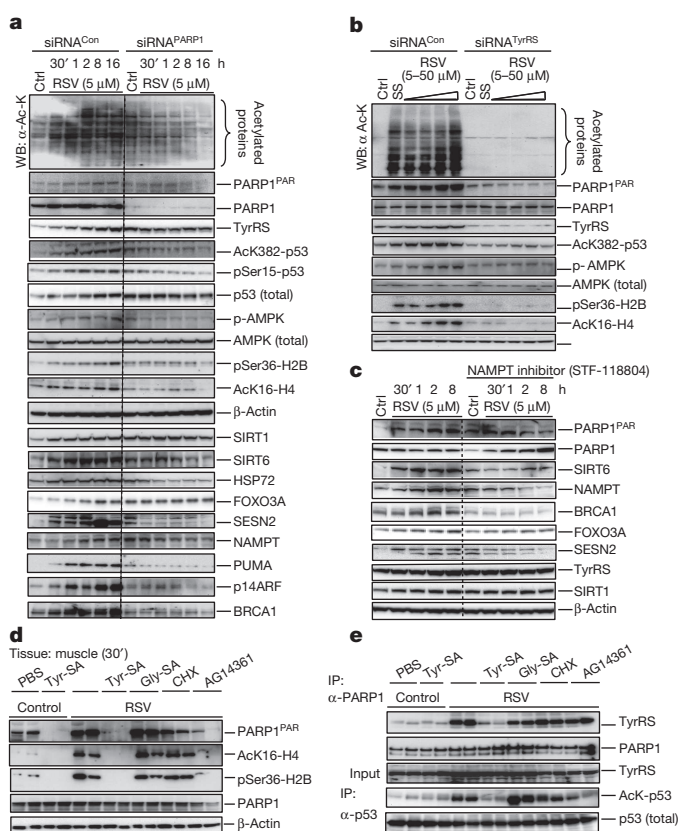


Figure 4 | Resveratrol treatment activates TyrRS–PARP1-driven signalling events in mouse tissues. **a**, **b**, siRNA of PARP1 (**a**) or siRNA of TyrRS (**b**) abrogates resveratrol-mediated downstream signalling events. HeLa cells were treated with siRNA^{PARP1} or siRNA^{TyrRS} for 60 h to knockdown PARP1 or TyrRS expression levels by approximately 70–80%. Knockdown efficiency was monitored by with α -PARP1 or α -TyrRS. For siRNA^{TyrRS}, HeLa cells were treated with either serum starvation or resveratrol (5, 10, 25, and 50 μM) for 45 min. The dashed line in Fig. 4a represents the demarcation between siRNA^{Con} and siRNA^{PARP1}; all the samples were run on the same gel. **c**, NAMPT inhibition abrogates resveratrol-mediated downstream signalling events. HeLa cells were pre-treated with STF-118804 (100 nM), a NAMPT inhibitor, for 16 h and then treated with resveratrol (5 μM), with samples collected at intervals as indicated. Various signalling events, as depicted, were monitored using appropriate antibodies. Acetylation status was monitored using α -acetyl-lysine. The dashed line separates results without and with the inhibitor. **d**, Resveratrol-mediated activation of PARP1 is blocked by Tyr-SA. Activations of Tip60 and AMPK were monitored by using α -AcK16-H4 and α -pSer36-H2B, respectively. **e**, Resveratrol-mediated interaction of TyrRS with PARP1 and acetylation of p53 are blocked by Tyr-SA. Immunoprecipitation of PARP1 and p53 from muscle tissue demonstrated RSV-mediated TyrRS–PARP1 interaction and p53 acetylation.

with AcK16-H4 and pSer36-H2B, were significantly increased in skeletal muscle (Extended Data Fig. 9a, b). Results in cardiac tissue were similar (Extended Data Fig. 9c, d). In addition, PARP1 activation diminished back to normal by 24 h (Extended Data Fig. 9e). Consistent with *ex vivo* assays, tissue samples from RSV-injected mice showed higher levels of TyrRS–PARP1 interaction together with increased auto-PARylation (Extended Data Fig. 9f).

Tyr-SA (5 μM) was added to the intravenous-injection with RSV. No interaction of TyrRS with PARP1 could be detected in skeletal or cardiac muscle and levels of PARylated and acetylated proteins did not increase. In contrast, co-injection of Gly-SA or cycloheximide (CHX, protein synthesis inhibitor) did not block RSV/TyrRS-mediated activation of PARP1 and its downstream signals (Fig. 4d (muscle) and Extended Data Fig. 9g, h (cardiac)). Consistent with *ex vivo* assays, immunoprecipitation of PARP1 (from harvested muscle tissue) pulled down TyrRS, and immunoprecipitation of p53 showed its increased

acetylation. Pull-down of TyrRS and p53 acetylation were blocked in mice that had co-injections of Tyr-SA or the AG14361 (SelleckChem) PARP1 inhibitor, but not of Gly-SA or cycloheximide (Fig. 4e). These responses *in vivo* parallel *ex vivo* assays, and support the idea that TyrRS is a major effector target for RSV that acts through PARP1 as described in Fig. 3a.

The mechanisms of action of RSV and of the stress response are both linked here to the activation of PARP1 through TyrRS and NAD⁺. The interaction of RSV with TyrRS could be viewed as an example of xenohormesis through interactions of a natural ligand with a protein target²⁹. In this instance, the natural ligand blocks the active site to create a tRNA synthetase catalytic null with a new, orthogonal function. This kind of catalytic null is in contradistinction to those created by alternative splicing events that specifically remove the active site⁷. The RSV–TyrRS activation of PARP1 is readily observable in a functional *in vitro* assay even at sub-micromolar concentrations (for example, 10 nM; Fig. 2a, middle). Thus, TyrRS–RSV-induced PARP1 activation appears at significantly lower RSV concentrations than seen with RSV functional binding to other targets^{1,25,30}. As a consequence, the direct effects of RSV binding to these targets of RSV are layered over a pre-existing foundation that comes from the TyrRS–RSV–PARP1–NAD⁺ connection.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 19 March; accepted 3 November 2014.

Published online 22 December 2014.

- Howitz, K. T. *et al.* Small molecule activators of sirtuins extend *Saccharomyces cerevisiae* lifespan. *Nature* **425**, 191–196 (2003).
- Viswanathan, M., Kim, S. K., Berdichevsky, A. & Guarente, L. A role for SIR-2.1 regulation of ER stress response genes in determining *C. elegans* life span. *Dev. Cell* **9**, 605–615 (2005).
- Baur, J. A. *et al.* Resveratrol improves health and survival of mice on a high-calorie diet. *Nature* **444**, 337–342 (2006).
- Milne, J. C. *et al.* Small molecule activators of SIRT1 as therapeutics for the treatment of type 2 diabetes. *Nature* **450**, 712–716 (2007).
- Jang, M. *et al.* Cancer chemopreventive activity of resveratrol, a natural product derived from grapes. *Science* **275**, 218–220 (1997).
- Fu, G., Xu, T., Shi, Y., Wei, N. & Yang, X. L. tRNA-controlled nuclear import of a human tRNA synthetase. *J. Biol. Chem.* **287**, 9330–9334 (2012).
- Lo, W. S. *et al.* Human tRNA synthetase catalytic nulls with diverse functions. *Science* **345**, 328–332 (2014).
- Wakasugi, K. & Schimmel, P. Two distinct cytokines released from a human aminoacyl-tRNA synthetase. *Science* **284**, 147–151 (1999).
- Guo, M. & Schimmel, P. Essential nontranslational functions of tRNA synthetases. *Nature Chem. Biol.* **9**, 145–153 (2013).
- Yang, X. L., Skene, R. J., McRee, D. E. & Schimmel, P. Crystal structure of a human aminoacyl-tRNA synthetase cytokine. *Proc. Natl Acad. Sci. USA* **99**, 15369–15374 (2002).
- Yang, X. L., Liu, F. M., Skene, R. J., McRee, D. E. & Schimmel, P. Crystal structure of an EMAP-II-like cytokine released from a human tRNA synthetase. *Helv. Chim. Acta* **86**, 1246–1257 (2003).
- Sajish, M. *et al.* Trp-tRNA synthetase bridges DNA-PKcs to PARP-1 to link IFN- γ and p53 signaling. *Nature Chem. Biol.* **8**, 547–554 (2012).
- Luo, X. & Kraus, W. L. On PAR with PARP: cellular stress signaling through poly(ADP-ribose) and PARP-1. *Genes Dev.* **26**, 417–432 (2012).
- Sinclair, D. A. & Guarente, L. Small-molecule allosteric activators of sirtuins. *Annu. Rev. Pharmacol. Toxicol.* **54**, 363–380 (2014).
- Kapoor, M., Otero, F. J., Slike, B. M., Ewalt, K. L. & Yang, X. L. Mutational separation of aminoacylation and cytokine activities of human tyrosyl-tRNA synthetase. *Chem. Biol.* **16**, 531–539 (2009).
- Lee, P. S., Zhang, H. M., Marshall, A. G., Yang, X. L. & Schimmel, P. Uncovering of a short internal peptide activates a tRNA synthetase procytokine. *J. Biol. Chem.* **287**, 20504–20508 (2012).
- Tang, Y., Zhao, W. H., Chen, Y., Zhao, Y. M. & Gu, W. Acetylation is indispensable for p53 activation. *Cell* **133**, 612–626 (2008).
- Bitterman, K. J., Anderson, R. M., Cohen, H. Y., Latorre-Esteves, M. & Sinclair, D. A. Inhibition of silencing and accelerated aging by nicotinamide, a putative negative regulator of yeast sir2 and human SIRT1. *J. Biol. Chem.* **277**, 45099–45107 (2002).
- Petes, S. J. & Lis, J. T. Activator-induced spread of poly(ADP-ribose) polymerase promotes nucleosome loss at Hsp70. *Mol. Cell* **45**, 64–74 (2012).
- Cohen-Armon, M. *et al.* DNA-independent PARP-1 activation by phosphorylated ERK2 increases Elk1 activity: a link to histone acetylation. *Mol. Cell* **25**, 297–308 (2007).
- Vyas, S. & Chang, P. Dual roles for PARP1 during heat shock: transcriptional activator and posttranscriptional inhibitor of gene expression. *Mol. Cell* **49**, 1–3 (2013).
- Hegan, D. C. *et al.* Inhibition of poly(ADP-ribose) polymerase down-regulates BRCA1 and RAD51 in a pathway mediated by E2F4 and p130. *Proc. Natl Acad. Sci. USA* **107**, 2201–2206 (2010).
- Orlando, G., Khoronenkova, S. V., Dianova, I. I., Parsons, J. L. & Dianov, G. L. ARF induction in response to DNA strand breaks is regulated by PARP1. *Nucleic Acids Res.* **42**, 2320–2329 (2014).
- Dasgupta, B. & Milbrandt, J. Resveratrol stimulates AMP kinase activity in neurons. *Proc. Natl Acad. Sci. USA* **104**, 7217–7222 (2007).
- Park, S. J. *et al.* Resveratrol ameliorates aging-related metabolic phenotypes by inhibiting cAMP phosphodiesterases. *Cell* **148**, 421–433 (2012).
- Butin-Israeli, V., Drayman, N. & Oppenheim, A. Simian virus 40 infection triggers a balanced network that includes apoptotic, survival, and stress pathways. *J. Virol.* **84**, 3431–3442 (2010).
- Bastide, B., Snoeckx, K. & Mounier, Y. ADP-ribose stimulates the calcium release channel RyR1 in skeletal muscle of rat. *Biochem. Biophys. Res. Commun.* **296**, 1267–1271 (2002).
- Shin, S. M., Cho, I. J. & Kim, S. G. Resveratrol protects mitochondria against oxidative stress through AMP-activated protein kinase-mediated glycogen synthase kinase-3 β inhibition downstream of poly(ADP-ribose)polymerase-LKB1 pathway. *Mol. Pharmacol.* **76**, 884–895 (2009).
- Howitz, K. T. & Sinclair, D. A. Xenohormesis: sensing the chemical cues of other species. *Cell* **133**, 387–391 (2008).
- Nwachukwu, J. C. *et al.* Resveratrol modulates the inflammatory response via an estrogen receptor-signal integration network. *Life* **3**, e02057 (2014).

Acknowledgements This work was supported in part the National Cancer Institute grant CA92577, by a fellowship from the National Foundation for Cancer Research, and by a Tyr Pharma through an agreement with The Scripps Research Institute. We thank the The Scripps Research Institute mouse facility for their efforts for this project. We also thank P. Chang for the ZZ-PARP1 clone, and the PARG and PARG-MT proteins, and Y. Shi for independently repeating some of the key experiments. We thank P. Chang, J. H. Chung, L. Guarente, L. Krauss, D. Sinclair, and C. Westphal for comments and suggestions on this work.

Author Contributions M.S. and P.S. conceived the idea, designed the research, analysed the data, and wrote the manuscript. M.S. performed the experiments.

Author Information X-ray structure coordinates of resveratrol-bound TyrRS and L-Tyr bound TyrRS have been deposited in Protein Data Bank under accession numbers 4Q93 and 4QBT, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to P.S. (schimmel@scripps.edu).

METHODS

Construction of the working model. Because both RSV and serum starvation activate PARP1 leading to the metabolic breakdown of NAD⁺ into nicotinamide and ADP-ribose, downstream signalling markers associated with the NAD⁺ metabolic flux were investigated. Targets were selected if they were activated and/or inhibited under conditions of RSV treatment, stress, or NAD⁺ metabolic by-product treatment (nicotinamide and ADP-ribose). For example, nicotinamide activates Tip60 (ref. 31) through SIRT1 inhibition¹⁸ and Tip60 activates ATM³², the p53 kinase. Protein p53 is a master regulator of NRF2 through p21 induction³³ and inducible expression of survival genes such as SIRT6 (ref. 34), SIRT1 (ref. 35), sestrins³⁶, and FOXO3A³⁷. (SIRT6 promotes DNA repair under stress by activating PARP1 (ref. 38) and SIRT1 downregulates the over-activation of PARP1 (ref. 39).) Similarly, ADP-ribose (metabolic by-product of NAD⁺ and an inhibitor of SIRT1⁴⁰) is a potent Ca²⁺ channel opener^{27,41}, a substrate for AMP production⁴², and hence an indispensable component for AMPK activation by resveratrol⁴³ and the stress response by phosphorylating H2B⁴⁴. AMPK is known to upregulate NAMPT expression⁴⁵, the major regulator of NAD⁺ levels in the cell⁴⁶. NAD⁺ is major substrate for sirtuins (deacetylation or mono-ADP-ribosylation^{38,47,48}) and PARPs ((mono/poly)-ADP-ribosylation^{13,49}). Although not depicted, p38 and NF- κ B (major transducers of stress signalling) are also regulated by TAK1 in a PARP1/ATM-dependent manner^{50,51}.

Antibodies. All antibodies were obtained from Cell Signaling Technology unless otherwise mentioned and those as follows: α -ATM was from Abcam and GeneTex; α -PARP1 and α -poly(ADP-ribose) were from BD Biosciences; α -p84 was from Sigma; α -pSer36-H2B from ECM Biosciences; rabbit α -human TyrRS polyclonal antibodies were homemade. α -Tip60 was from Millipore; α -Ack16-H4 was from Active Motif; and α -Hsp72 (catalogue number SPA-812) was from Assay Designs.

Cell culture, transfection, and resveratrol treatment. HeLa cells (from ATCC, mycoplasma free) were cultured in a humidified incubator with 5% CO₂ in DMEM medium (Invitrogen) supplemented with 10% FBS (Invitrogen) and 1 \times penicillin/streptomycin. The cells were transfected with pcDNA6-TyrRS-V5 (wild type), or pcDNA3.1-ZZ-PARP1 using Lipofectamine LTX (Invitrogen). Resveratrol (Santa-Cruz Biotechnologies), at a series of concentrations from 0 to 50 μ M, was used to treat HeLa cells 0–24 h in experiments described. As the effects of RSV are strongly influenced by the background level of PARP1 activity, the experiments were done only after checking to verify that background PARP1 activation was low. Our criterion was to see PARP1 activation upon serum starvation (DMEM medium without FBS), along with Hsp72 induction⁵². When serum starvation induced PARP1 activation with Hsp72 induction within 30–45 min (from the same batch of a split of HeLa cells), we consistently saw PARP1 activation at lower concentrations of RSV (1 μ M). Cell culture medium was supplemented with Tyr-SA or Gly-SA (RNA-TEC) where mentioned. Silencer Select Pre-Designed siRNA against PARP1 (siRNA s1099, 5'-GCAGCUUCAUAACCGAAGAtt-3'), SIRT1 (siRNA s23770, 5'-GGCUUG AUGGUAUCAGUAtt-3') and TyrRS (siRNA s443, 5'-GGACUUUGCUCG GAGGUAtt-3') were purchased from Invitrogen and transfected into HeLa cells using Lipofectamine RNAiMAX (Invitrogen). NAMPT, the rate-limiting enzyme in the NAD⁺ salvage pathway, was inhibited using STF-118804⁵³ according to the manufacturer's protocol (SelleckChem). The effect of a concomitant increase in nicotinamide production with a reduction of NAD⁺ concentration (leading to SIRT1 inhibition) was monitored by immunoblotting with α -Ack382-p53, an acetylation site known to be specifically targeted by SIRT1 (ref. 1). Monitoring of α -pSer 15-p53 indicated if enhanced acetylation preceded the phosphorylation event⁵⁴. Total p53 was blotted with α -p53. Transactivation of p53 was determined by monitoring its known targets, such as p21, PUMA, 14-3-3, FOXO3A, SESN2, and SIRT6. Activation of NRF2 was further confirmed by following HO-1 expression, using the cognate antibodies. Activation of AMPK (Thr 172-phosphorylation) mediated through AMP and Ca²⁺ influx²⁵ was monitored using α -pThr 172-AMPK. Activation of AMPK on its targets was further determined by both α -pSer 36-H2B and by the expression level of NAMPT.

Cell viability assay. HeLa (1 \times 10⁶) cells were reverse-transfected with siRNAs and viability was monitored using an RTCA iCELLigence System (ACEA Biosciences). HeLa cells were cultured in a humidified incubator with 5% CO₂ in DMEM medium (Invitrogen) supplemented with 10% FBS (Invitrogen) and 1 \times penicillin/streptomycin.

Preparation of cell fractions. Protocol for cell fractionation was followed as described previously¹². Briefly, for whole-cell lysate preparation, HeLa cells were dissolved in 1 \times SDS-polyacrylamide gel electrophoresis (SDS-PAGE) loading buffer containing 300 mM NaCl. For cytoplasmic fraction preparation, HeLa cells were suspended in 0.1 ml of swelling buffer for 6 min on ice, then incubated with 0.1 ml of plasma membrane lysis buffer for 5 min on ice. The cells were immediately passed through a 21-gauge needle ten times and centrifuged for 10 min at 3220g, 4 °C. The supernatant was harvested, while the pellets were used for the preparation of nuclear fractions. For this purpose, the pellet (nuclei) was incubated for 30 min at 4 °C with 0.2 ml nuclear extraction buffer. Whenever required, PARP1 (AG14361) and

PARG inhibitors (ADP-HPD, Millipore) were added to the cell lysis buffer to ensure unwanted PARylation and removal of PAR chains with PARG.

Immunoprecipitation and Ni-NTA pull-down assays. Protocol for immunoprecipitation was followed as described previously¹². Briefly, the supernatants were pre-cleared by incubation with protein G beads. The pre-cleared cell lysates were incubated at 4 °C for 1 h with either α -PARP1, α -ATM, α -DNA-PKcs, α -TyrRS, α -Tip60, or non-immune immunoglobulin-G (IgG), at a concentration of 5 μ g ml⁻¹ followed by incubation with 30 μ l Protein G beads (pretreated with 10 mg ml⁻¹ BSA) at 4 °C for 1 h with rotation. PARP1 (AG14361) and PARG inhibitors (ADP-HPD, Millipore) were added to the cell lysis buffer to ensure unwanted PARylation and removal of PAR chains with PARG. Immunoprecipitates were washed three times, subjected to SDS-PAGE and immunoblotted with specific antibodies. Whenever mentioned, the ZZ domain allowed immunoprecipitation of ectopically expressed ZZ-PARP1, using anti-IgG. Whenever Ni-NTA pull-down was performed, proteins with a 6 \times -His tag were overexpressed in *Escherichia coli*. Cells were lysed and the supernatant fractions containing the soluble proteins were mixed with HeLa cell lysates. For Ni-NTA pull-downs, normal procedures for immunoprecipitation were followed with 15–20 mM imidazole in the washing buffer.

Animal experiments. All experiments were approved by The Scripps Research Institute Institutional Animal Care and Use Committee (protocol number 13-0002) and conducted by the mice facility at The Scripps Research Institute. BALB/cByJ mice were originally purchased from Jackson Laboratories. Six-week-old male mice were kept with a 12 h light–dark cycle with free access to food and water for 3 days before conducting the experiment. For all studies, mice were dosed once and collected sample after either at 30 min or at 24 h. Briefly, activation of PARP1 in mouse tissues treated with resveratrol was performed by the intravenous injection of a 100 μ l sample of resveratrol (10 μ M) in PBS into eight mouse tails (0.012 mg kg⁻¹). A 100 μ l PBS injection was used as a control in six mice. Tissue samples were collected from four resveratrol-treated mice and three control mice after 30 min treatment. Tissue samples from the remaining mice were collected after 24 h. In a different experiment, blocking of resveratrol-mediated activation of PARP1 by Tyr-SA was analysed with five groups of two mice each, a 100 μ l sample of resveratrol (10 μ M) in PBS alone and, separately, with Tyr-SA (5 μ M), Gly-SA (5 μ M), cycloheximide (5 μ M), or the AG14361 PARP1 inhibitor (10 μ M) was injected intravenously. A 100 μ l PBS injection with Tyr-SA was used in a control group (two mice) as indicated. Tissue samples were collected after 30 min treatment. Muscle or heart tissues samples were homogenized and analysed by immunoblotting for the status of poly-ADP-ribosylation and associated signalling events.

Recombinant protein purification. Human TyrRS, PARP1 and their variants were purified as previously described^{10,12}. Briefly, DNA encoding either full-length TyrRS (amino acids 1–528), mini-TyrRS (amino acids 1–341), dN-TyrRS (amino acids 237–528), or CT-TyrRS (amino acids 324–528) was cloned into Nde-I/Hind-III sites of pET-20b vector (Novagen). The expressed proteins included a 6-His tag from the vector sequence. The full-length PARP1 and its variants (amino (N)- and C-terminal domains) were cloned into pET 20b vector. All proteins having a C-terminal His-tag were expressed in *E. coli* strain BL21 (DE3) by induction for 4 h with 1 mM isopropyl β -D-thiogalactopyranoside. Proteins were purified from the supernatants of lysed cells using Ni-NTA agarose (Qiagen) column chromatography according to the manufacturer's instructions. We also included an additional high-salt two-column wash (1–1.5 M NaCl), to remove the endogenous DNA/RNA associated with the purified protein. The NaCl concentration was increased to 1.5 M in increments and returned to final elution buffer having 250 mM NaCl in decrements. All the proteins were subjected to a gel-filtration (S-200) chromatography and the protein peak corresponding to homogenous protein was collected. As endotoxin associated with purified protein interfered with the inhibitory/activating effect of RSV on TyrRS/PARP1, the complete protein purification included an additional endotoxin removal step. The purified protein was further passed through a column containing Detoxi-GelTM (Pierce-Thermo Scientific) and followed the manufacturer's instructions. The final endotoxin level (0.5 EU mg⁻¹ ml⁻¹) was measured using an Endosafe^R-PTS kit (Charles River Laboratories). The quality of each protein purification was validated by SDS-PAGE analysis.

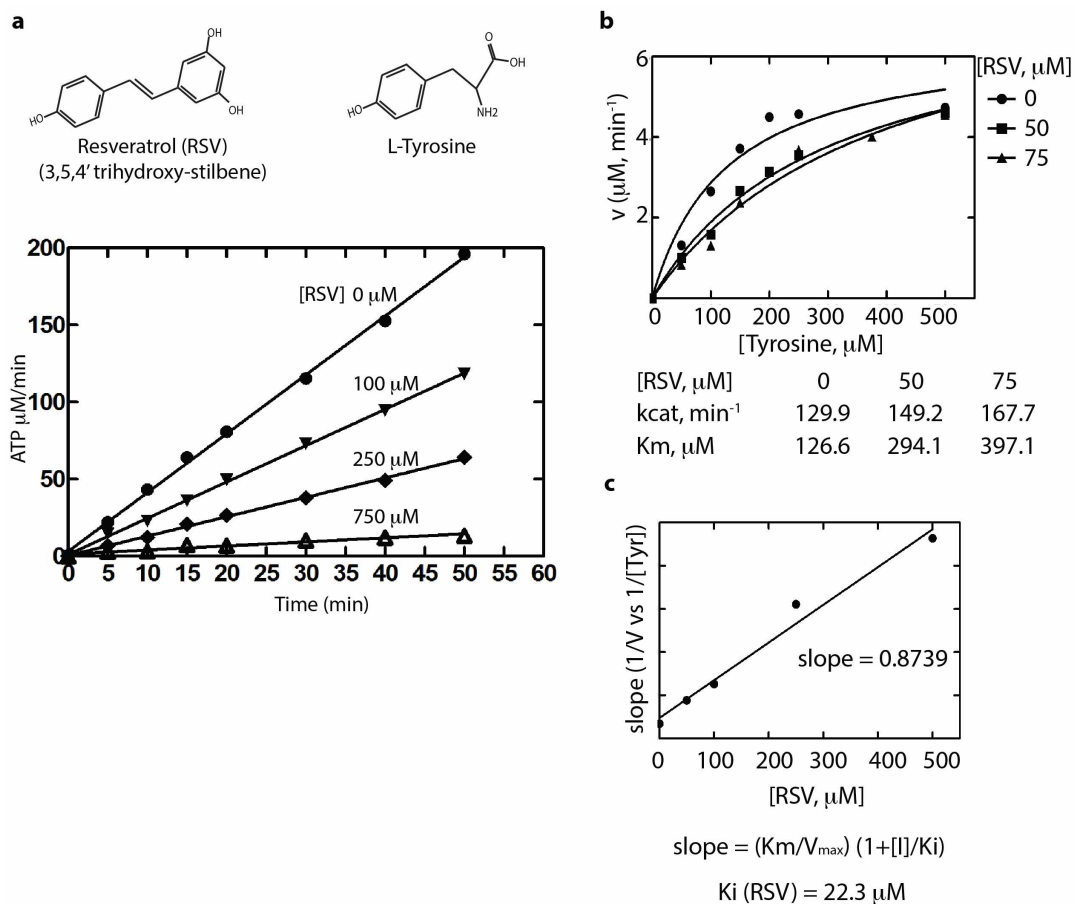
Protein crystallization and data collection. All the steps were followed as mentioned in the previous report for mini-TyrRS crystallization¹⁰. Briefly, mini-TyrRS was mixed with either 1 mM RSV or 2 mM tyrosine and incubated at 4 °C for 16 h. Before setting up the crystallization trials, the protein samples were subjected to high-speed centrifugation (15,871 relative centrifugal force for 15 min) to remove all the precipitates and the clear soluble fraction transferred to a new tube. The crystallization trials were done using 2.1 M (NH₄)₂SO₄, 0.1 M NaH₂PO₄/K₂HPO₄ (pH 6–8), and 2% acetone at room temperature (22–25 °C). Although, mini-TyrRS crystals with L-tyrosine were grown in 3 days (data collected at Stanford Synchrotron Radiation Lightsource), crystals with RSV were grown only after 3–4 months at around pH 7. X-ray data were collected at 2.1 Å using the in-house X-ray diffraction facility (The Scripps Research Institute). Data were integrated and scaled using

HKL2000. The electron density was refined by molecular replacement (CNS and CCP4/REFMAC suite and Coot) using the known structure of mini-TyrRS (Extended Data Table 1) and structures were deposited in PDB (accession numbers 4Q93 and 4QBT).

ATP-pyrophosphate exchange assay. Tyrosyl adenylate synthesis was measured by using the tyrosine-dependent ATP-pyrophosphate (PPi) exchange assay. A mixture containing 100 mM HEPES (pH 7.5), 20 mM KCl, 2 mM ATP, 1 mM NaPPi, 2 mM DTT, 250–500 μ M L-tyrosine, 10 mM MgCl_2 , and about 0.01 mCi ml^{-1} Na^{32}P PPi was added to 50–100 nM purified TyrRS (endotoxin free, 0.5 EU mg^{-1} ml^{-1}), pre-incubated with 0–1 mM RSV at 4 °C for 30 min. The ATP-PPi exchange reaction was incubated at room temperature (22–25 °C), and aliquots were removed at specified time intervals and quenched in a mixture containing 40 mM NaPPi, 1.4% HClO_4 , 0.4% HCl, and 8% (w/v) of activated charcoal. After thoroughly mixing, the charcoal was filtered and washed with a solution of 7% HClO_4 and 200 mM NaPPi using Spin-X Centrifuge Filters (Corning) containing 0.45- μ m pore-size cellulose acetate filters. After drying, the charcoal was punched into scintillation vials and the radioactivity of the ATP bound to the charcoal mixture was measured by scintillation counting.

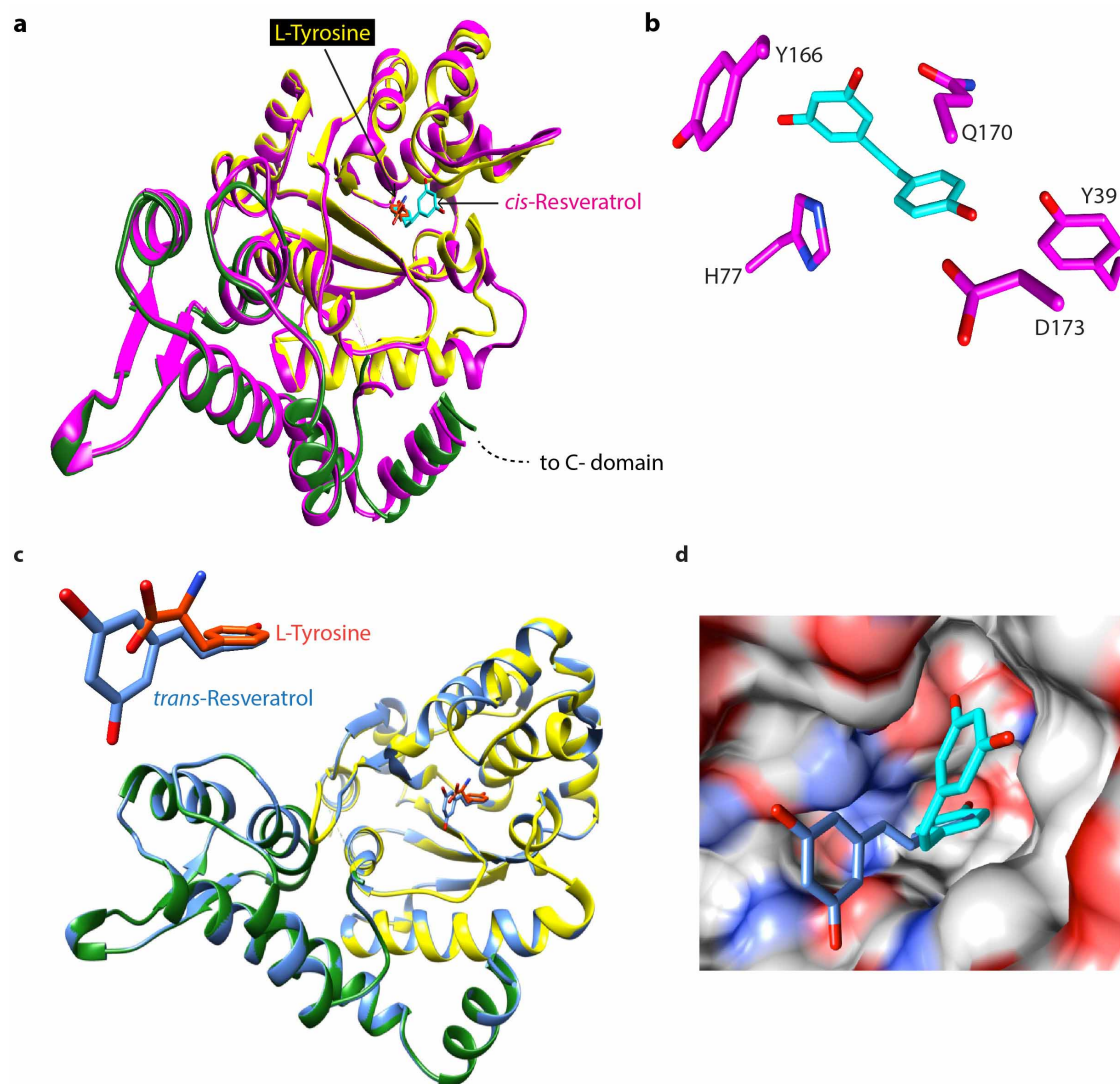
In vitro PARylation assay. Protocol was as previously described¹². Briefly, 5 μ l recombinant PARP1 (2 μ g) was mixed with 5 μ l recombinant TyrRS (0–50 μ M) and, depending on the experiment, either RSV, TyrSA, DNA, or AG14361 were also added (5 μ l) and adjusted the volume to 25 μ l. This mixture was incubated at 4 °C for 15 min, then incubated with 25 μ l 2 \times PAR assay cocktail (50 mM HEPES, pH 7.5, 100 mM KCl, 10 mM MgCl_2 , 0.2 mM EGTA, 0.1 mM EDTA, 40 nM NAD, and 1 μ Ci ^{32}P NAD (Perkin-Elmer)) at 30 °C for 30 min. The RSV concentration used in Fig. 2a (middle) was 0, 5, 10, 20, 30, 40, 60, and 80 nM, and in Fig. 2a (bottom) was 40 nM. The reaction was stopped by addition of 50 μ l 2 \times SDS sample buffer (5% SDS) and heated at 95 °C for 1 min. The ^{32}P -PARylated PARP1 was separated from free ^{32}P NAD on SDS-PAGE, vacuum-dried, and the poly-ADP-ribosylation activity of PARP1 was analysed using a phosphorimager (Typhoon FLA 7000, GE Healthcare).

31. Wang, J. D. & Chen, J. J. SIRT1 regulates autoacetylation and histone acetyltransferase activity of TIP60. *J. Biol. Chem.* **285**, 11458–11464 (2010).
32. Sun, Y., Jiang, X., Chen, S., Fernandes, N. & Price, B. D. A role for the Tip60 histone acetyltransferase in the acetylation and activation of ATM. *Proc. Natl Acad. Sci. USA* **102**, 13182–13187 (2005).
33. Chen, W. M. *et al.* Direct interaction between Nrf2 and p21(Cip1/WAF1) upregulates the Nrf2-mediated antioxidant response. *Mol. Cell* **34**, 663–673 (2009).
34. Zhang, P. *et al.* Tumor suppressor p53 cooperates with SIRT6 to regulate gluconeogenesis by promoting FoxO1 nuclear exclusion. *Proc. Natl Acad. Sci. USA* **111**, 10684–10689 (2014).
35. Nemoto, S., Fergusson, M. M. & Finkel, T. Nutrient availability regulates SIRT1 through a forkhead-dependent pathway. *Science* **306**, 2105–2108 (2004).
36. Budanov, A. V. & Karin, M. p53 target genes sestrin1 and sestrin2 connect genotoxic stress and mTOR signaling. *Cell* **134**, 451–460 (2008).
37. Renault, V. M. *et al.* The pro-longevity gene FoxO3 is a direct target of the p53 tumor suppressor. *Oncogene* **30**, 3207–3221 (2011).
38. Mao, Z. *et al.* SIRT6 promotes DNA repair under stress by activating PARP1. *Science* **332**, 1443–1446 (2011).
39. Rajamohan, S. B. *et al.* SIRT1 promotes cell survival under stress by deacetylation-dependent deactivation of poly(ADP-ribose) polymerase 1. *Mol. Cell. Biol.* **29**, 4116–4129 (2009).
40. Smith, B. C., Hallows, W. C. & Denu, J. M. A continuous microplate assay for sirtuins and nicotinamide-producing enzymes. *Anal. Biochem.* **394**, 101–109 (2009).
41. Perraud, A. L. *et al.* ADP-ribose gating of the calcium-permeable LTRPC2 channel revealed by Nudix motif homology. *Nature* **411**, 595–599 (2001).
42. Formentini, L. *et al.* Poly(ADP-ribose) catabolism triggers AMP-dependent mitochondrial energy failure. *J. Biol. Chem.* **284**, 17668–17676 (2009).
43. Hawley, S. A. *et al.* Use of cells expressing γ subunit variants to identify diverse mechanisms of AMPK activation. *Cell Metab.* **11**, 554–565 (2010).
44. Bungard, D. *et al.* Signaling kinase AMPK activates stress-promoted transcription via histone H2B phosphorylation. *Science* **329**, 1201–1205 (2010).
45. Fulco, M. *et al.* Glucose restriction inhibits skeletal myoblast differentiation by activating SIRT1 through AMPK-mediated regulation of Nampt. *Dev. Cell* **14**, 661–673 (2008).
46. Ramsey, K. M. *et al.* Circadian clock feedback cycle through NAMPT-mediated NAD⁺ biosynthesis. *Science* **324**, 651–654 (2009).
47. Haigis, M. C. & Guarente, L. P. Mammalian sirtuins—emerging roles in physiology, aging, and calorie restriction. *Genes Dev.* **20**, 2913–2921 (2006).
48. Michishita, E. *et al.* SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin. *Nature* **452**, 492–496 (2008).
49. Vyas, S. *et al.* Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nat. Commun.* **5**, 4426 (2014).
50. Yang, Y. *et al.* A cytosolic ATM/NEMO/RIP1 complex recruits TAK1 to mediate the NF- κ B and p38 mitogen-activated protein kinase (MAPK)/MAPK-activated protein 2 responses to DNA damage. *Mol. Cell. Biol.* **31**, 2774–2786 (2011).
51. Stilmann, M. *et al.* A nuclear poly(ADP-ribose)-dependent signalosome confers DNA damage-induced I κ B kinase activation. *Mol. Cell* **36**, 365–378 (2009).
52. Martin, N. *et al.* PARP-1 transcriptional activity is regulated by sumoylation upon heat shock. *EMBO J.* **28**, 3534–3548 (2009).
53. Matheny, C. J. *et al.* Next-generation NAMPT inhibitors identified by sequential high-throughput phenotypic chemical and functional genomic screens. *Chem. Biol.* **20**, 1352–1363 (2013).
54. Lambert, P. F., Kashanchi, F., Radonovich, M. F., Shiekhata, R. & Brady, J. N. Phosphorylation of p53 serine 15 increases interaction with CBP. *J. Biol. Chem.* **273**, 33048–33053 (1998).



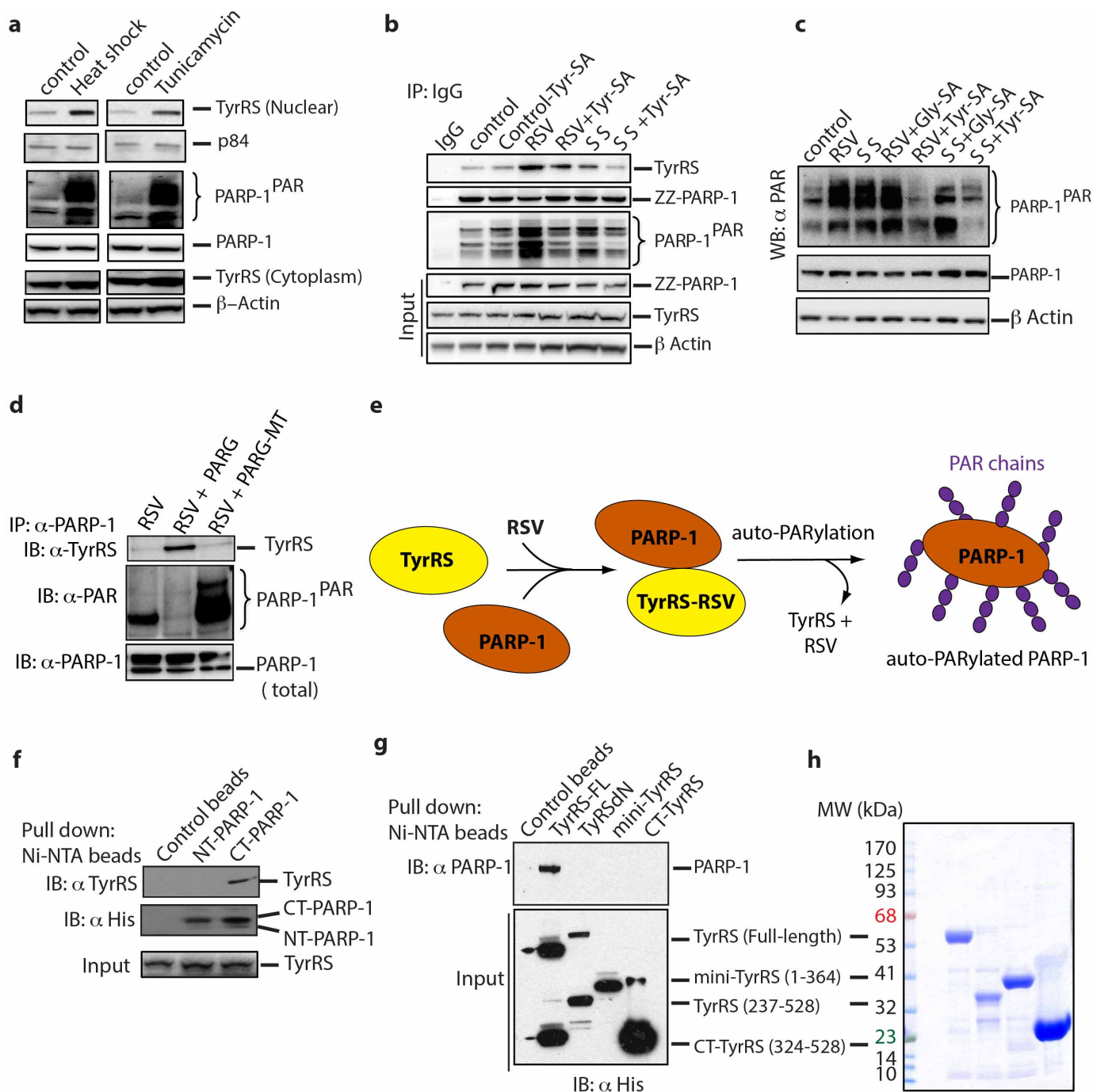
Extended Data Figure 1 | Resveratrol inhibits TyrRS activation. **a**, The ATP-P_i exchange assay as described in Methods demonstrated the inhibitory effect of resveratrol on TyrRS and **b** resveratrol shifts the Michaelis constant

(K_{m}) for tyrosine. **c**, Resveratrol binds TyrRS better than tyrosine. The apparent K_{i} for resveratrol was deduced by varying the concentration of RSV and plotting the slope of ($1/v$ versus $1/[\text{Tyr}]$) versus [RSV] as indicated.



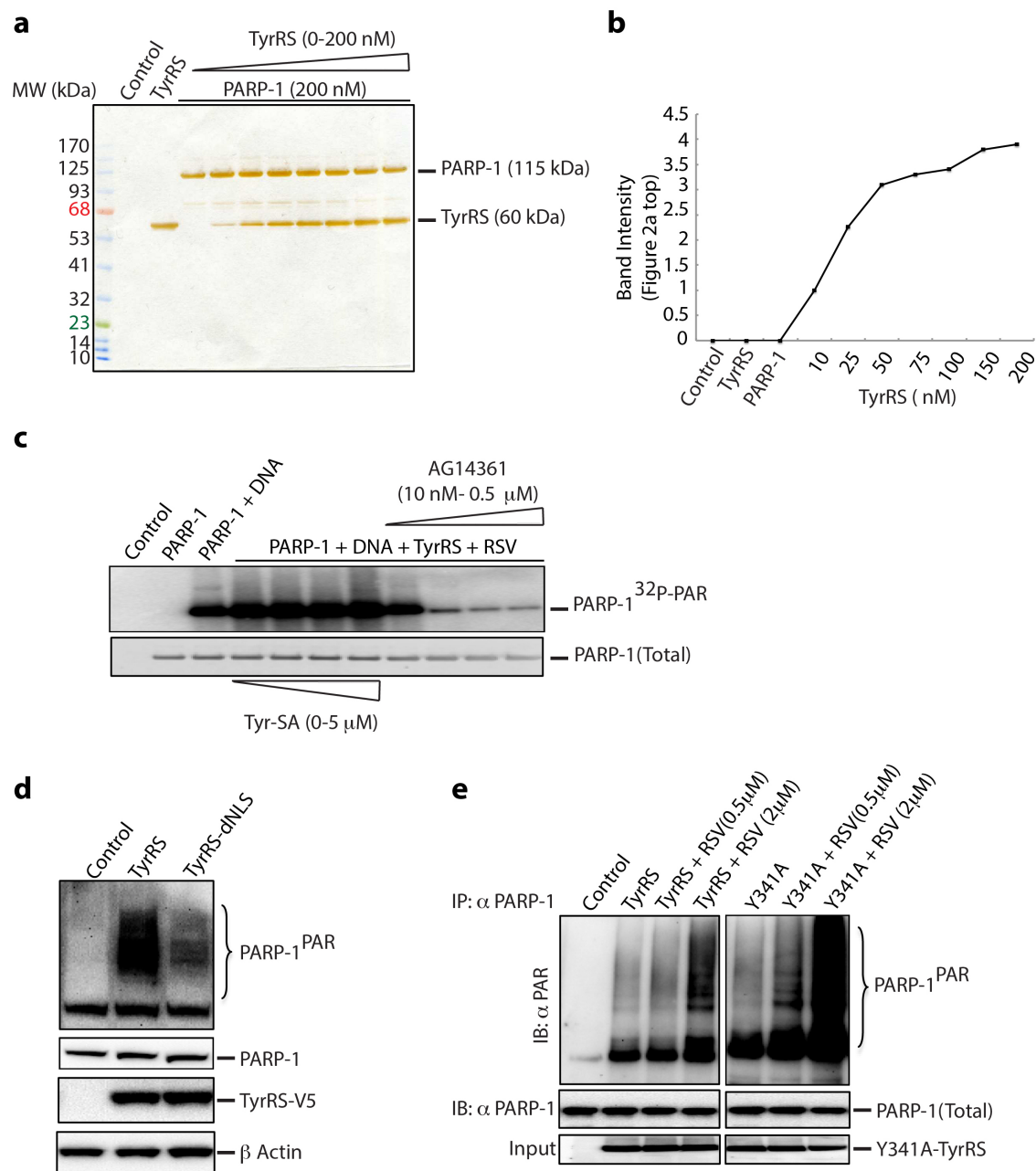
Extended Data Figure 2 | Resveratrol induces a distinct conformational change upon binding to active site of TyrRS. **a**, Comparison of the overall conformational change induced by resveratrol at the active site of TyrRS by structure-based superposition (yellow, tyrosine-bound structure; magenta, resveratrol-bound structure). Note the conformational change near the helix region (P331–P342) that connects the linker region with the C-domain. **b**, Illustration of the extensive interactions of resveratrol with the active site.

c, The *trans*-resveratrol (dark blue) docks (manual) into TyrRS active site without significant structural disturbances. **d**, Generation of a new pocket through a RSV-induced conformational change in TyrRS accommodates the dihydroxy phenolic ring of RSV (otherwise exposed to the destabilizing aqueous environment in the *trans* form) and hence facilitates the *trans* (dark blue) to *cis* (light blue) conversion of RSV.



Extended Data Figure 3 | Resveratrol facilitates the TyrRS-PARP1 interaction in an active-site-dependent manner. **a**, Both heat shock (42 °C for 30 min) and tunicamycin-treatment (10 μg ml⁻¹, endoplasmic reticulum stress) facilitated the nuclear translocation of TyrRS and activation of PARP1. **b**, Resveratrol or serum starvation facilitate TyrRS interaction with PARP1, and Tyr-SA prevents this interaction. ZZ-PARP1 was immunoprecipitated with IgG from HeLa cells treated with RSV or serum starvation alone or in combination with Tyr-SA. **c**, Resveratrol- or serum-starvation-mediated PARP1 activation is blocked only by Tyr-SA and not by Gly-SA. **d**, TyrRS interacts directly with PARP1. HeLa cell lysate after RSV treatment (5 μM,

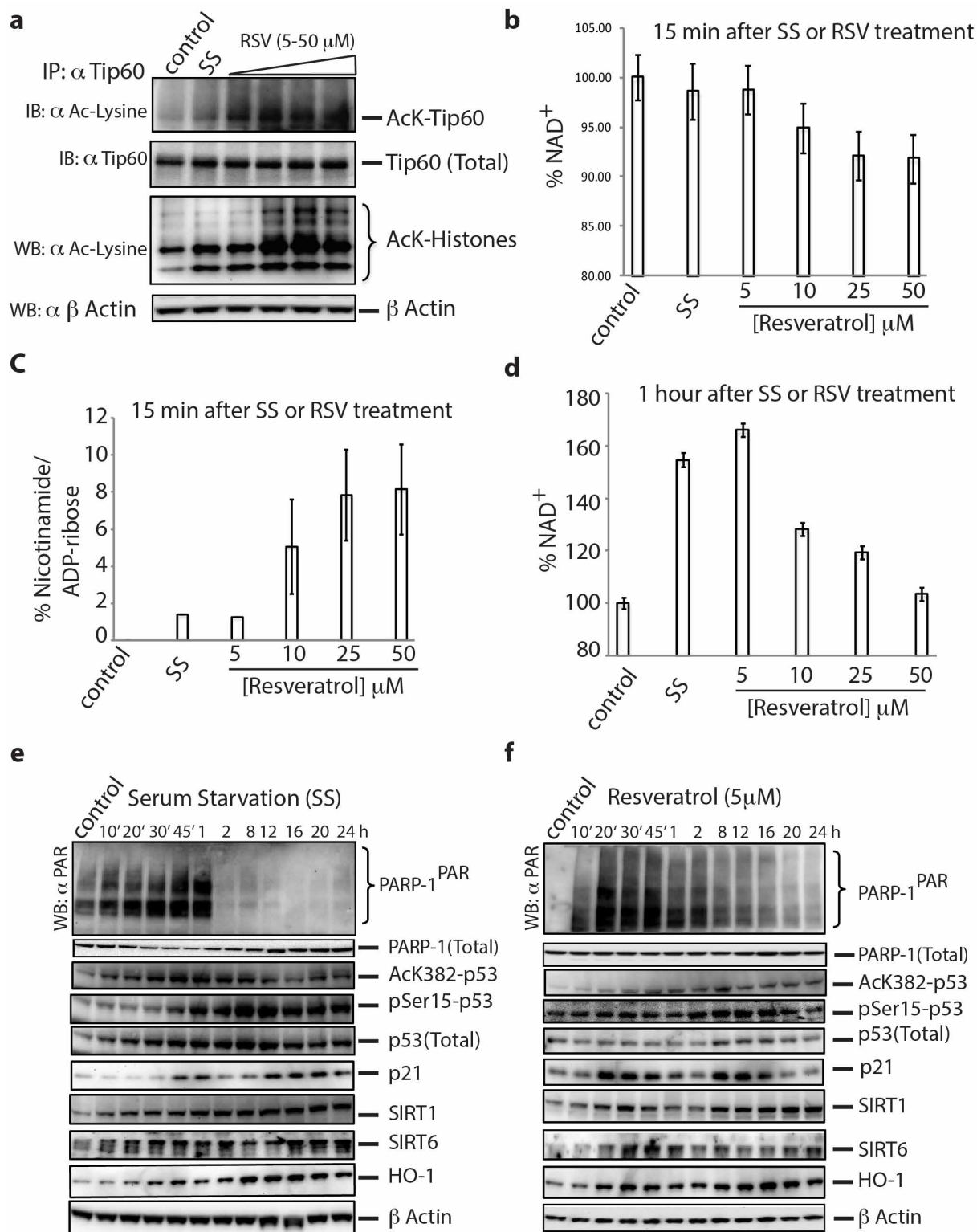
30 min) was divided into three parts and treated with PARG and catalytically inactive PARG-MT. PARP1 was immunoprecipitated and analysed for TyrRS interaction. **e**, Model illustrating the mechanism of RSV-mediated TyrRS interaction with PARP1 and subsequent release after auto-PARYlation. **f**, Ni-NTA pull-down of N- and C-terminal fragments of PARP1 overexpressed in *E. coli* demonstrates that TyrRS interacts with the C-terminal region of PARP1. **g**, Only the full-length TyrRS (1-528), but none of the various fragments of TyrRS (mini-TyrRS (1-364), ΔN-TyrRS (228-528), or the C-domain (328-528)), interacts with PARP1. **h**, Coomassie blue staining of a gel showing the total protein input in for the experiment of Extended Data Fig. 3g.



Extended Data Figure 4 | Tyrosyl-AMP analogue (Tyr-SA) does not affect DNA-dependent auto-PARYlation of PARP1. **a**, Silver-stained SDS-PAGE gel showing the purity and input of PARP1 and TyrRS in the *in vitro* PARYlation study of Fig. 2. **b**, Quantitation (Image J software) of the band intensity of PARYlated PARP1 in Fig. 2a, top. **c**, Tyrosyl-AMP analogue

(Tyr-SA) does not affect DNA-dependent auto-PARYlation of PARP1.

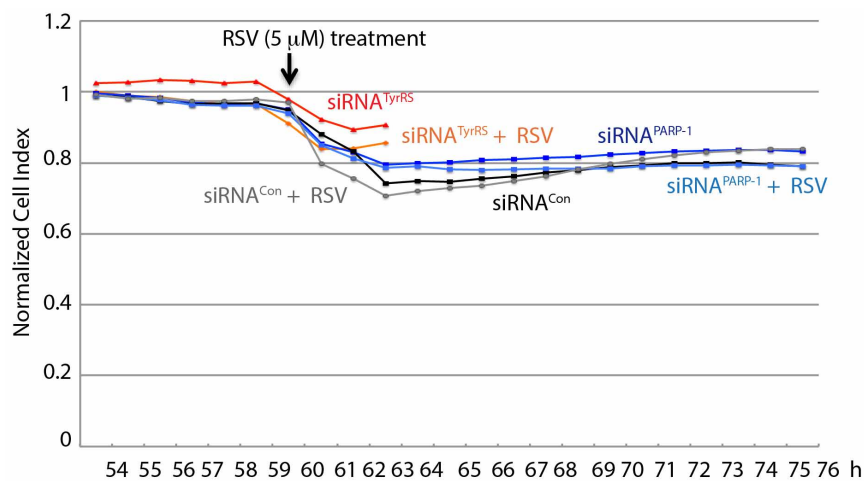
d, Overexpression of nuclear-translocation-weakened mutant of TyrRS⁶ is less effective in activating PARP1. **e**, Y314A-TyrRS is more sensitive to RSV than is TyrRS in facilitating PARP1 activation.



Extended Data Figure 5 | Resveratrol enhances the acetylation of Tip60 and modulates NAD⁺ concentration in a dose- and time-dependent manner.

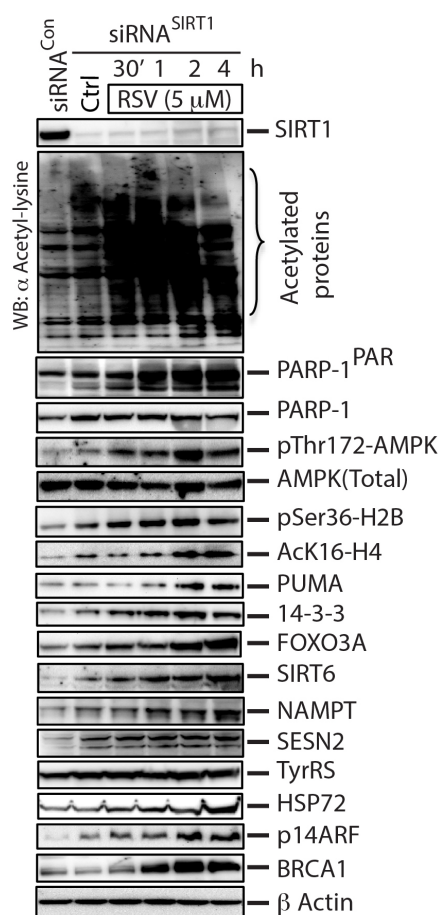
a, Treatment of HeLa cells (1 h) with increasing concentration of resveratrol enhances the acetylation level of Tip60. Activation of Tip60 was monitored by histone acetylation status. **b**, Total NAD⁺ contents of serum-starved cells or RSV-treated samples were compared with untreated samples at 15 min using a commercially available BioVision NAD⁺/NADH quantitation colorimetric kit. **c**, Total nicotinamide or ADP-ribose produced was deduced from the difference in the amount of NAD⁺ in each sample with respect to the untreated sample (consumption of one mole of NAD⁺ would give rise one mole of nicotinamide and one mole of ADP-ribose). **d**, Total NAD⁺ content of the

serum-starved cells or RSV-treated samples were compared with untreated samples at 1 h. (Although the experiments were done in biological triplicates (all samples showing similar results), the error bars in the figure represent the deviations from the mean of the technical triplicates from one representative biological sample.) **e**, **f**, Time course study of poly-ADP-ribosylation status and associated signalling events after (e) serum starvation (extended time course data of the same image shown in Fig. 3c) and (f) treatment with 5 μ M RSV. Using the respective antibodies, Activation of p53 was monitored by the induction of p21 and SIRT6. Activation of NRF2 was monitored by HO-1 induction.

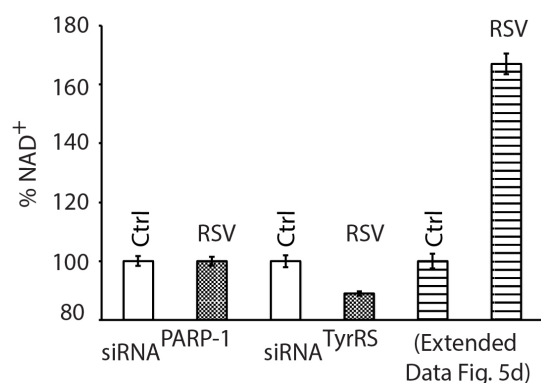


Extended Data Figure 6 | siRNA (siRNA^{TyrRS} or siRNA^{PARP1}), with and without low RSV (5 μ M), does not affect cell viability. HeLa cells (1×10^6) were reverse-transfected with siRNA targeted to TyrRS or PARP1. An siRNA^{Con} (a scrambled sequence of siRNA^{PARP1}) was used as a control. Viability was monitored using the RTCA iCELLigence System (ACEA

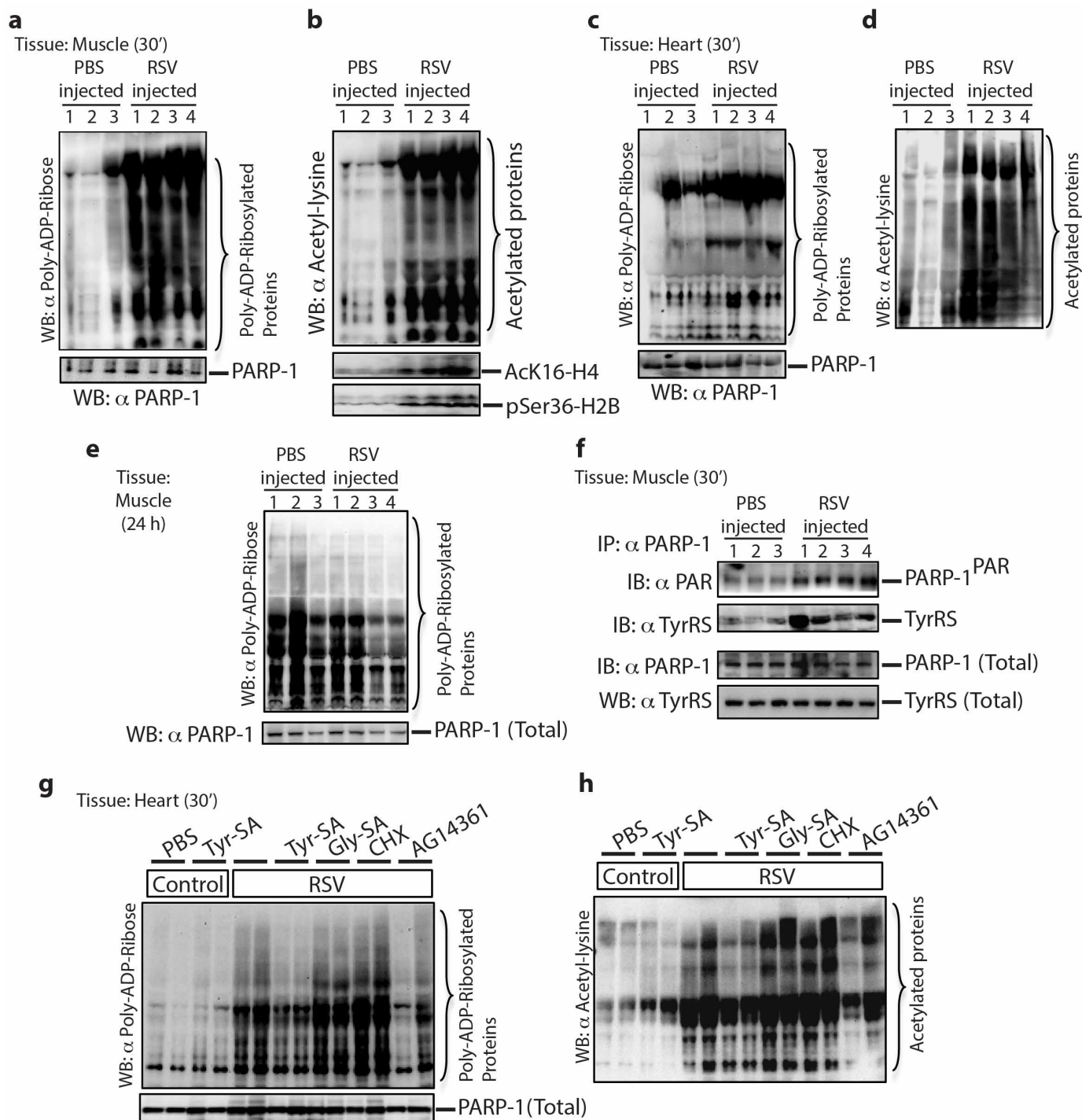
Biosciences). Samples were treated with RSV (5 μ M) at 60 h and monitoring was continued for another 2 h for siRNA^{TyrRS} (a total of 62 h of monitoring) and for another 16 h for siRNA^{Con} and siRNA^{PARP1} (total 76 h monitoring).



Extended Data Figure 7 | siRNA^{SIRT1} did not affect downstream signalling events at low RSV (5 μM). HeLa cells were treated with siRNA^{SIRT1} for 60 h to knockdown SIRT1. HeLa cells were treated with RSV (5 μM) for another 4 h and samples were collected intervals as indicated. Samples were analysed for downstream signalling markers using appropriate antibodies.



Extended Data Figure 8 | siRNA- (siRNA^{TyrRS} or siRNA^{PARP1}) treated cells did not upregulate the levels of NAD⁺ in response to RSV (5 μ M) after 1 h. HeLa cells (1×10^6) were reverse transfected, separately, with siRNA targeted to PARP1 or TyrRS. A scrambled sequence of target siRNA was used as a control. The total NAD⁺ content of RSV (5 μ M)-treated samples was compared with untreated samples at 1 h, using a commercially available BioVision NAD⁺/NADH quantitation colorimetric kit. (Although the experiments were done in biological triplicates (all samples showing similar results), the error bars in the figure represent the deviations from the mean of the technical triplicates from one representative biological sample.) The comparator (shown as a dashed bar) is taken from Extended Data Fig. 5d.



Extended Data Figure 9 | Resveratrol treatment activates PARP1 and associated signalling events in the mouse tissues. **a**, Activation of PARP1 in mouse muscle tissue treated with resveratrol monitored by increased PARylation and **b** by increased acetylation status. Activations of Tip60 and AMPK were monitored by using α -AcK16-H4 and α -pSer36-H2B, respectively. **c**, Activation of PARP1 in mouse heart tissue treated with resveratrol monitored by increased PARylation and **d** by increased acetylation

status. **e**, Resveratrol treatment causes only a transient activation on PARP1. Immunoblotting of mouse muscle tissue samples after 24 h of RSV treatment showed no significant difference in the level of PARP1^{PAR} compared with control. **f**, RSV treatment enhances TyrRS interaction with and activation of PARP1 in the muscle tissue. **g** and **h**, Resveratrol-mediated activation of PARP1 (**g**, monitored by PARylation status; **h**, monitored by acetylation status) is blocked by Tyr-SA in mouse heart tissues.

Extended Data Table 1 | Data collection and refinement statistics

	Resveratrol-TyrRS	L-Tyrosine-TyrRS
Data collection		
Space group	P21212	P21212
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	73.97, 162.55, 35.0	73.97, 162.55, 35.0
α , β , γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	37 - 2.1(2.15-2.09)*	34.63 - 2.1(2.15-2.09)*
<i>R</i> _{sym} or <i>R</i> _{merge}	27.1	44.7
<i>I</i> / σ <i>I</i>	4.1	4.6
Completeness (%)	98.6	98
Redundancy	3.2	5.2
Refinement		
Resolution (Å)	2.1	2.1
No. reflections	25438	24513
<i>R</i> _{work} / <i>R</i> _{free}	0.19/0.24	0.22/0.27
No. atoms	2955	2858
Protein	2653	2636
Ligand/ion	16	12
Water	284	238
B-factors		
Protein	37.664	38.199
Ligand/ion	43.2	42.3
Water	47.3	48.4
R.m.s deviations		
Bond lengths (Å)	0.018	0.017
Bond angles (°)	2.017	2.008

*Highest resolution shell is shown in parenthesis.

Numerical parameters for the RSV-TyrRS and L-Tyr-TyrRS co-crystals are listed and compared side-by-side in separate columns.

NAD captureSeq indicates NAD as a bacterial cap for a subset of regulatory RNAs

Hana Cahová^{1*}, Marie-Luise Winz^{1*}, Katharina Höfer^{1*}, Gabriele Nübel¹ & Andres Jäschke¹

A distinctive feature of prokaryotic gene expression is the absence of 5'-capped RNA. In eukaryotes, 5',5'-triphosphate-linked 7-methylguanosine protects messenger RNA from degradation and modulates maturation, localization and translation¹. Recently, the cofactor nicotinamide adenine dinucleotide (NAD) was reported as a covalent modification of bacterial RNA². Given the central role of NAD in redox biochemistry, posttranslational protein modification and signalling^{3,4}, its attachment to RNA indicates that there are unknown functions of RNA in these processes and undiscovered pathways in RNA metabolism and regulation. The unknown identity of NAD-modified RNAs has so far precluded functional analyses. Here we identify NAD-linked RNAs from bacteria by chemo-enzymatic capture and next-generation sequencing (NAD captureSeq). Among those identified, specific regulatory small RNAs (sRNAs) and sRNA-like 5'-terminal fragments of certain mRNAs are particularly abundant. Analogous to a eukaryotic cap, 5'-NAD modification is shown *in vitro* to stabilize RNA against 5'-processing by the RNA-pyrophosphohydrolase RppH⁵ and against endonucleolytic cleavage by ribonuclease (RNase) E⁶. The nudix phosphohydrolase NudC⁷ decaps NAD-RNA and thereby triggers RNase-E-mediated RNA decay, while being inactive against triphosphate-RNA. *In vivo*, ~13% of the abundant sRNA RNAI is NAD-capped in the presence, and ~26% in the absence, of functional NudC. To our knowledge, this is the first description of a cap-like structure and a decapping machinery in bacteria.

To isolate NAD-RNA from biological samples, a new capturing technique was required, which had to be specific for the NAD moiety without damaging the RNA molecules. NAD glycohydrolases appeared suitable but NADase from porcine brain contained RNases, and recombinant human CD38 was not capable of preparative transglycosylation on NAD. Adenosine diphosphate-ribosylcyclase (ADPRC) from *Aplysia californica*⁸, previously described to replace nicotinamide with other N-heterocyclic nucleophiles⁹, was surprisingly found to be capable of catalysing the transglycosylation reaction of NAD with alkynyl alcohols. ADPRC requires a labile N-glycosidic bond, as in the positively charged pyridine, but not in canonical purine and pyrimidine nucleosides¹⁰. The 'clickable' transglycosylation product was biotinylated by a copper-catalysed azide-alkyne cycloaddition^{11,12}, for which the overall yield was highest using 4-pentyn-1-ol, both with free NAD (Extended Data Fig. 1) and with *in vitro* transcribed 5'-NAD-RNA¹³. No reaction product was detected with NADH and canonical RNA. After capturing the biotin-linked RNAs on streptavidin beads, they were selectively amplified by a protocol combining 3'-adaptor ligation, reverse transcription, tailing of complementary DNA, ligation of a second adaptor and polymerase chain reaction (PCR). We applied this workflow (NAD captureSeq; Fig. 1) to total RNA from either a modified *Escherichia coli* JM109 laboratory strain or a more wild-type-like *E. coli* K-12 strain to construct three different next-generation sequencing (NGS) libraries (Extended Data Fig. 2a–c).

In all three data sets (Supplementary Data 1–3), comparison of reads from samples (B), which were fully treated, versus controls (A), which

lacked ADPRC, allowed a clear identification of enriched RNAs among the most abundant sequences (Fig. 2 and Extended Data Figs 2d–i, 3). Many enriched candidates were sRNAs reported to act by different pathways^{14,15}, for example, GadY, GcvB, ChiX, McaS, RNAI and CopA (Fig. 2 and Extended Data Table 1). The latter two were observed only in JM109, as they are replication control RNAs encoded on plasmids absent in our K-12 strain^{16,17}. Other sRNAs were found not to be enriched. Another set of enriched sequences represented 5'-fragments of different mRNAs, for example, of genes *gatY*, *pgk*, *hdeD*, and of the leader peptide genes *ilvL* and *hisL* (Extended Data Fig. 3). Besides the leader peptides with regulatory functions, several of the encoded enzymes are involved in cellular metabolism (for example, PGK¹⁸). Other enzymes and several enriched sRNAs act in stress responses (for example, response to acid, enzyme HdeD¹⁹, and sRNAs GadY¹⁵, GcvB²⁰ and ChiX²¹, or response to ultraviolet stress, enzyme UspE²²). In contrast, no ribosomal RNA or transfer RNA was enriched. Apart from annotated transcription starts at adenosine for most enriched genes, no common sequence or structural features were identified for the enriched RNAs in promoters, transcripts, or reported transcription and σ factors.

To verify that the enrichment observed in NAD captureSeq represents a difference in cDNA amounts rather than PCR bias during NGS library preparation, we performed real-time PCR with cDNA collected after the second adaptor ligation, before PCR. As reference RNAs, we chose species that were detected but not enriched in NGS (Extended Data Fig. 4a). These data supported our results: after normalization to tRNA^{Set} and 5S rRNA, RNAI and CopA RNA were enriched >500-fold, while four other sRNAs yielded 24- to 163-fold enrichments. To confirm that—despite the lack of a fragmentation step in NAD captureSeq—specific 5'-terminal mRNA fragments, rather than full-length mRNAs were enriched, real-time PCRs were performed with primers targeting different loci within the transcript. Up to ~400-fold enrichment of mRNA 5'-termini was corroborated, whereas primers targeting the middle of those mRNAs yielded no or decreased enrichment (Extended

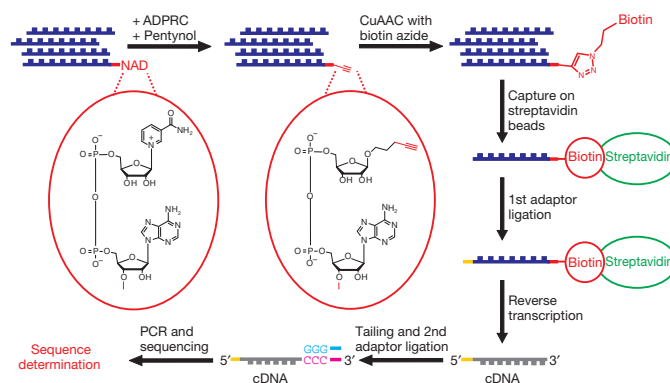


Figure 1 | NAD captureSeq protocol applied to total RNA. CuAAC, copper-catalysed azide-alkyne cycloaddition; pentynol, 4-pentyn-1-ol.

¹Institute of Pharmacy and Molecular Biotechnology (IPMB), Heidelberg University, 69120 Heidelberg, Germany.

*These authors contributed equally to this work.

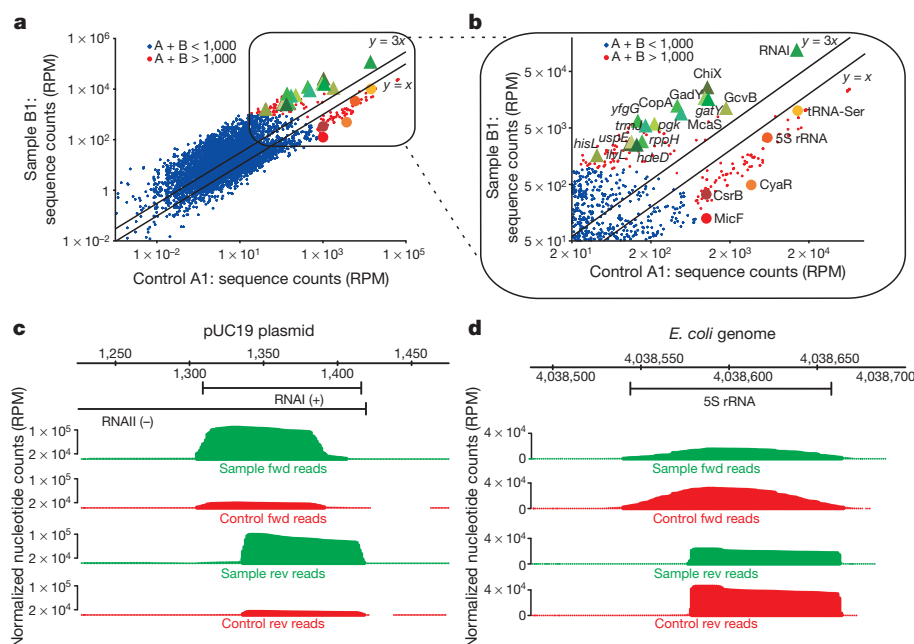


Figure 2 | Analysis of enriched RNAs by NGS. **a, b,** Abundance of RNAs in fully treated sample (B1) versus minus ADPRC control (A1) found by NGS analysis of JM109 samples (120–160 bp) (reads per million (RPM), average of forward and reverse read counts). Triangles, enriched; circles, non-enriched.

Initial capitals indicate sRNAs. **c,** NGS analysis of RNAi sequences found in controls (A1, red) and samples (B1, green) in forward (fwd) and reverse (rev) reads. **d,** NGS analysis of 5S rRNA sequences.

Data Fig. 4b, d). Additionally, northern blot analysis of non-treated total RNA (Extended Data Fig. 4c) confirmed the presence of short 5'-terminal fragments, indicating that they are not generated during NAD captureSeq, but rather represent aborted or processed transcripts or hitherto unknown sRNAs. These 5'-fragments exist in an NAD-modified state and may have additional (for example, regulatory) roles.

To prove that RNAs identified by NAD captureSeq do indeed bear the NAD modification and have not been isolated for other reasons, highly abundant RNAi^{17,23} and 5S rRNA were pulled-down from total RNA, nuclease P1 digested and analysed by high performance liquid chromatography mass spectrometry (HPLC-MS)^{2,24}. As expected, the extracted ion chromatogram for the species with the mass to charge ratio (m/z) = 540.05 ($[M-H]^+$) corresponding to the main NAD fragment formed during MS ionization², revealed a characteristic peak for *in vitro* transcribed model NAD-RNA. This peak was also detected for cellular RNAi, but not for the 5S rRNA control (Fig. 3a, b), for which similar amounts of the canonical nucleotides, but no NAD or NAD fragments, were measured. The peak disappeared upon treatment with ADPRC and 4-pentyn-1-ol before digestion, while a new peak (m/z = 624.10) appeared, representing the reaction product. For the less abundant sRNA GcvB, the HPLC-MS protocol was modified to detect the $[M]^+$ ion of NAD, m/z = 664.116, which was—again—not present in the 5S rRNA control (Extended Data Fig. 5).

The 5'-modification status of RNA (triphosphate versus monophosphate) was recently reported to determine the stability and processing of RNAs themselves⁶ and of sRNA targets²⁵. We speculated that the NAD modification may act in a similar way to the eukaryotic cap and stabilize regulatory RNAs. Indeed, *in vitro* transcribed NAD-RNAi is strongly stabilized (18.5-fold compared with 5'-monophosphate-RNAi and 3.5-fold compared with 5'-triphosphate-RNAi; Fig. 3c–e) against RNase E, the enzyme that initiates many decay pathways in *E. coli*.

RNA pyrophosphohydrolase (RppH) has been described to convert 5'-triphosphate-RNA to 5'-monophosphate-RNA, and thereby trigger RNase-E-mediated decay⁵. We observe that NAD modification strongly (60-fold) decelerates RNAi processing by RppH, while a less pronounced effect is observed for NADH (sixfold; Fig. 3f, g). RppH is one of 13 nudix phosphohydrolases^{26,27} described in *E. coli*. Another member of this class is NudC, which hydrolyses NAD(H) into NMN(H) and AMP⁷. We

find that this enzyme is highly active in decapping 5'-NAD(H)-RNAi but is virtually inactive against 5'-triphosphate-RNAi (Fig. 3h, i). A NudC mutant in which a critical residue in the nudix motif is exchanged (NudC(E178Q)) was completely inactive against transcribed 5'-NAD-RNAi (Extended Data Fig. 6a, b). 5'-NAD-RNAi treated with NudC for 20 min is subsequently processed by RNase E rapidly and at a rate comparable with 5'-triphosphate-RNAi treated with RppH, while the other enzyme–substrate combinations result in much slower degradation (Extended Data Fig. 6c, d, compare with Fig. 3d, e).

To examine the biological significance of the NAD cap and to substantiate the enzymatic decapping activity, we subjected cellular total RNA to NudC decapping followed by splinted ligation *in vitro* (Extended Data Fig. 7a). We found that RNAi was decapped by wild-type NudC, but not by the E178Q mutant, in a time-dependent manner (Extended Data Fig. 7b, c). These findings provide further support for the presence of NAD-capped RNA, and for the 5'-terminus as the attachment site² *in vivo*. Using this knowledge, we determined the percentage of NAD modification for RNAi by NudC decapping of dephosphorylated total RNA and specific ligation of NudC-decapped RNAi sequences (Fig. 4a, b and Extended Data Fig. 7d–h). This analysis revealed that ~13% of full-length RNAi contained the NAD cap, whereas 6S RNA, which was not enriched in NAD captureSeq, tested negative. Upon chromosomal *nudC* deletion, the NAD modification level of full-length RNAi was doubled to ~26%. Complementation of this defect *in trans* by a plasmid-borne copy of wild-type *nudC* but not by an inactive mutant (*nudC* E178Q) reduced the NAD modification back to wild-type level. Despite these differences in NAD modification state between wild type and the $\Delta nudC$ knockout strain, no effect on the half-life of RNAi or GcvB RNA was measured *in vivo* when transcription was blocked by addition of the antibiotic rifampicin (Extended Data Fig. 6e–h). These data suggest that further, currently unknown, players are involved in 5'-NAD-RNA processing, or that only a subset of NAD-RNAs is processed by NudC *in vivo*. Our observations are strikingly similar to a recent report on 5'-triphosphate processing of sRNAs, in which knockout of RppH, the enzyme generally accepted to perform 5'-triphosphate removal, did not result in major changes in 5'-modification status²⁵. These combined findings highlight the complexity of RNA processing, which is still poorly understood even in the model organism *E. coli*.

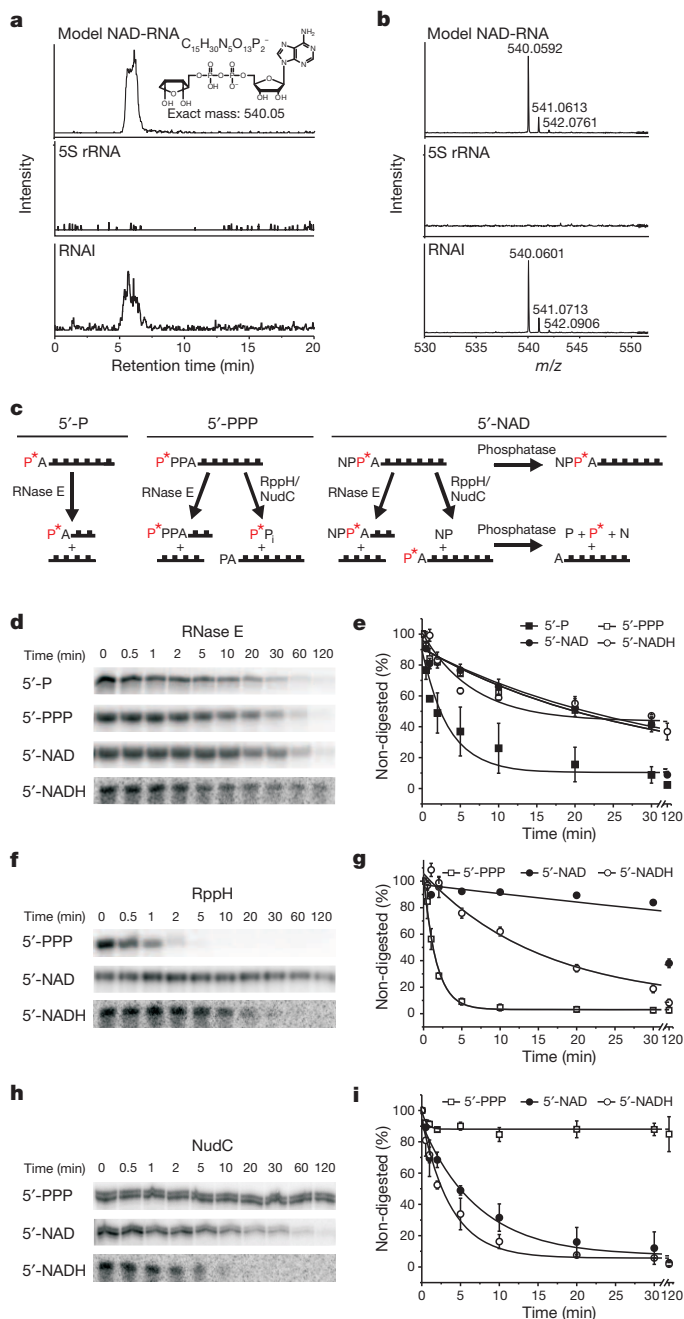


Figure 3 | Characterization of NAD-RNAI. **a**, HPLC-MS extracted ion chromatogram ($m/z = 540.05$) of nuclease-P1-digested model NAD-RNA, 5S rRNA and RNAI. **b**, High-resolution MS of HPLC fractions eluting at 5–6 min in **a**. **c**, Schematic representation of RNase E degradation and RppH/NudC cleavage of *in vitro* transcribed RNAI with different radiolabelled (red asterisks) 5'-termini (P, phosphate; NPPA = NAD). **d–i**, 5'-Dependent RNase E degradation (**d**, **e**), RppH processing (**f**, **g**) and NudC decapping (**h**, **i**) of RNAI. In **d**, radioactive full-length bands disappear due to cleavage. In **f**, **h**, radiolabels are lost due to 5'-end processing and 5'-NAD(H) samples contained alkaline phosphatase to remove ^{32}P -phosphate accessible after RppH/NudC processing. Data points in **e**, **g**, and **i** represent mean \pm standard deviation (s.d.), $n = 2$.

To our knowledge, this work represents the first demonstration of a cap-like structure in prokaryotes, which modulates the susceptibility of specific RNAs towards RNase E and phosphohydrolases. The decapping activity of NudC could provide the bacterium with a mechanism to selectively initiate degradation for a subset of cellular RNAs orthogonal to RppH processing of 5'-triphosphate-RNA (Fig. 4c). In eukaryotes,

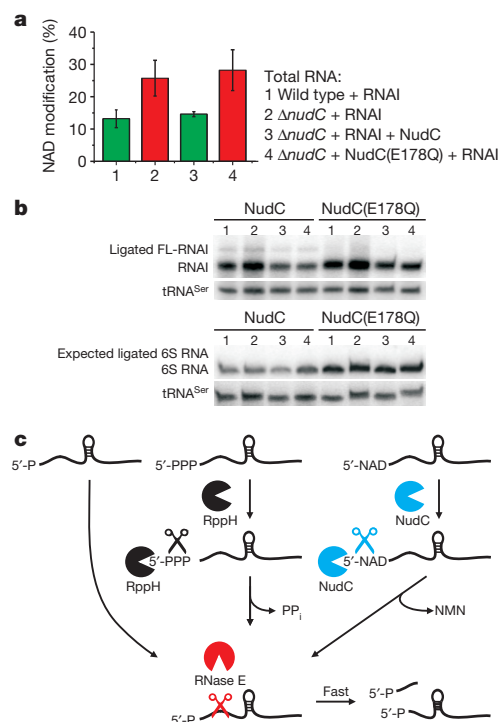


Figure 4 | Quantification of NAD modification of RNAI- and NAD-decapping activity of NudC in *E. coli*. **a**, 5'-Modification state of full-length RNAI (FL-RNAI) *in vivo* in the presence or absence of functional NudC (see Extended Data Fig. 7 for details). For each strain: $n = 3$ (independent cultures); mean \pm s.d. **b**, Exemplary northern blots of RNAI and 6S RNA for the quantification of NAD modification. 6S RNA (not enriched in NAD captureSeq (Extended Data Table 1)) tested negative. No ligation product was obtained upon treatment of dephosphorylated total RNA with dysfunctional NudC(E178Q). tRNA^{Ser} served as loading control. **c**, Alternative RNase E-dependent RNA degradation pathways. PP_i, pyrophosphate; NMN, nicotinamide mononucleotide.

nudix phosphohydrolases are known to act as decapping enzymes, but such activity was hitherto unknown in prokaryotes²⁷. Since 5' modification with monophosphate was recently shown to modulate RNase E-mediated degradation of sRNA targets²⁵, it is possible that the NAD cap influences not only the turnover, but also the function, of NAD-RNAs.

In addition to stabilization and specific degradation, NAD capping could serve to recruit specific proteins to the RNA 5'-end, as NAD is among the most common protein ligands, with well-characterized binding pockets in many proteins²⁸. This could endow the RNAs or associated molecular machines (for example, the degradosome²⁹) with additional structural, catalytic or regulatory properties. As many NAD-binding proteins distinguish between reduced and oxidized NAD³⁰, associated proteins may differ depending on redox state. Besides the identification of further players in NAD-RNA processing, future work will focus on the mechanism of NAD capping and different potential functions of the NAD cap to reveal the scope of NAD capping in prokaryotic biology.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 16 July 2013; accepted 27 October 2014.

Published online 22 December 2014.

1. Topisirovic, I., Svitkin, Y. V., Sonenberg, N. & Shatkin, A. J. Cap and cap-binding proteins in the control of gene expression. *RNA* **2**, 277–298 (2011).
2. Chen, Y. G., Kowtoniuk, W. E., Agarwal, I., Shen, Y. & Liu, D. R. LC/MS analysis of cellular RNA reveals NAD-linked RNA. *Nature Chem. Biol.* **5**, 879–881 (2009).
3. Houtkooper, R. H., Cantó, C., Wanders, R. J. & Auwerx, J. The secret life of NAD⁺: an old metabolite controlling new metabolic signaling pathways. *Endocr. Rev.* **31**, 194–223 (2010).

4. Oppenheimer, N. J. NAD⁺ and NADP⁺ as prosthetic groups for enzymes. *eLS* <http://dx.doi.org/10.1002/9780470015902.a0000637.pub2> (2010).
5. Deana, A., Celesnik, H. & Belasco, J. G. The bacterial enzyme RppH triggers messenger RNA degradation by 5' pyrophosphate removal. *Nature* **451**, 355–358 (2008).
6. Mackie, G. A. Ribonuclease E is a 5'-end-dependent endonuclease. *Nature* **395**, 720–724 (1998).
7. Frick, D. N. & Bessman, M. J. Cloning, purification, and properties of a novel NADH pyrophosphatase. Evidence for a nucleotide pyrophosphatase catalytic domain in MutT-like enzymes. *J. Biol. Chem.* **270**, 1529–1534 (1995).
8. Walseth, T. F. & Lee, H. C. Synthesis and characterization of antagonists of cyclic-ADP-ribose-induced Ca²⁺ release. *Biochim. Biophys. Acta* **1178**, 235–242 (1993).
9. Preugschat, F., Tomberlin, G. H. & Porter, D. J. The base exchange reaction of NAD⁺ glycohydrolase: identification of novel heterocyclic alternative substrates. *Arch. Biochem. Biophys.* **479**, 114–120 (2008).
10. Migaud, M. E., Pederick, R. L., Bailey, V. C. & Potter, B. V. Probing *Aplysia californica* adenosine 5'-diphosphate ribosyl cyclase for substrate binding requirements: design of potent inhibitors. *Biochemistry* **38**, 9105–9114 (1999).
11. Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Edn Engl.* **41**, 2596–2599 (2002).
12. Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regioselective copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
13. Huang, F. Efficient incorporation of CoA, NAD and FAD into RNA by *in vitro* transcription. *Nucleic Acids Res.* **31**, e8 (2003).
14. Opdyke, J. A., Fozo, E. M., Hemm, M. R. & Storz, G. RNase III participates in GadY-dependent cleavage of the *gadX-gadW* mRNA. *J. Mol. Biol.* **406**, 29–43 (2011).
15. Opdyke, J. A., Kang, J. G. & Storz, G. GadY, a small-RNA regulator of acid response genes in *Escherichia coli*. *J. Bacteriol.* **186**, 6698–6705 (2004).
16. Gerhart, E., Wagner, H. & Nordström, K. Structural analysis of an RNA molecule involved in replication control of plasmid R1. *Nucleic Acids Res.* **14**, 2523–2538 (1986).
17. Lacatena, R. M. & Cesareni, G. Base pairing of RNA I with its complementary sequence in the primer precursor inhibits ColE1 replication. *Nature* **294**, 623–626 (1981).
18. D'Alessio, G. & Josse, J. Glyceraldehyde phosphate dehydrogenase, phosphoglycerate kinase, and phosphoglyceromutase of *Escherichia coli*. Simultaneous purification and physical properties. *J. Biol. Chem.* **246**, 4319–4325 (1971).
19. Masuda, N. & Church, G. M. Regulatory network of acid resistance genes in *Escherichia coli*. *Mol. Microbiol.* **48**, 699–712 (2003).
20. Jin, Y., Watt, R. M., Danchin, A. & Huang, J. D. Small noncoding RNA GcvB is a novel regulator of acid resistance in *Escherichia coli*. *BMC Genomics* **10**, 165 (2009).
21. Hayes, E. T. *et al.* Oxygen limitation modulates pH regulation of catabolism and hydrogenases, multidrug transporters, and envelope composition in *Escherichia coli* K-12. *BMC Microbiol.* **6**, 89 (2006).
22. Gustavsson, N., Diez, A. & Nyström, T. The universal stress protein paralogues of *Escherichia coli* are co-ordinately regulated and co-operate in the defence against DNA damage. *Mol. Microbiol.* **43**, 107–117 (2002).
23. Eguchi, Y. & Tomizawa, J. Complex formed by complementary RNA stem-loops and its stabilization by a protein: function of ColE1 Rom protein. *Cell* **60**, 199–209 (1990).
24. Dumelin, C. E., Chen, Y., Leconte, A. M., Chen, Y. G. & Liu, D. R. Discovery and biological characterization of geranylated RNA in bacteria. *Nature Chem. Biol.* **8**, 913–919 (2012).
25. Bandrya, K. J. *et al.* The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. *Mol. Cell* **47**, 943–953 (2012).
26. McLennan, A. G. Substrate ambiguity among the nudix hydrolases: biologically significant, evolutionary remnant, or both? *Cell. Mol. Life Sci.* **70**, 373–385 (2013).
27. Song, M. G., Bail, S. & Kiledjian, M. Multiple Nudix family proteins possess mRNA decapping activity. *RNA* **19**, 390–399 (2013).
28. Gherardini, P. F., Ausiello, G., Russell, R. B. & Helmer-Citterich, M. Modular architecture of nucleotide-binding pockets. *Nucleic Acids Res.* **38**, 3809–3816 (2010).
29. Bandrya, K. J., Bouvier, M., Carpousis, A. J. & Luisi, B. F. The social fabric of the RNA degradosome. *Biochim. Biophys. Acta* **1829**, 514–522 (2013).
30. Fjeld, C. C., Birdsong, W. T. & Goodman, R. H. Differential binding of NAD⁺ and NADH allows the transcriptional corepressor carboxyl-terminal binding protein to serve as a metabolic sensor. *Proc. Natl Acad. Sci. USA* **100**, 9202–9207 (2003).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the CellNetworks Deep Sequencing Core Facility for Solexa sequencing, M. Brunner for access to the light cycler, M. Helm for advice on RNA mass spectrometry and B. Luisi for the RNase E expression vector, as well as A. Krause, J. Becker, A. Samanta, M. Tesch, F. Siebert, L. Obenauer and other members of the Jäschke laboratory for help and discussions. H.C. was supported by a postdoctoral fellowship from the Alexander-von-Humboldt Foundation. M.-L.W. acknowledges a PhD fellowship from the Hartmut Hoffmann-Berling International Graduate School of Molecular & Cellular Biology. A.J. is supported by the Deutsche Forschungsgemeinschaft, SFB 623, the Federal Ministry of Education and Research (BMBF), and the Helmholtz Initiative on Synthetic Biology.

Author Contributions All authors designed the experiments, analysed and interpreted results and wrote the paper. H.C., M.-L.W., K.H. and G.N. performed the experiments.

Author Information NGS data have been deposited at <http://www.geneprof.org> under accession numbers gpXP_001108, gpXP_001123 and gpXP_001153. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.J. (jaeschke@uni-hd.de).

METHODS

General. Reagents were purchased from Sigma-Aldrich and used without further purification. Oligonucleotides were purchased from Biomers.net GmbH and Integrated DNA Technologies. MS experiments were performed on a Bruker micro-TOFQII ESI mass spectrometer. Recorded MS spectra were deconvoluted using maximum entropy deconvolution. For high-resolution mass spectra, internal calibration was performed (enhanced quadratic mode) using ESI Tune-mix (Fluka) as calibrant. Reversed-phase HPLC purification was performed on an Agilent 1100 Series HPLC system equipped with a diode array detector using Phenomenex Luna 5 μm C18 110 A (250 \times 4.60 mm) at a flow rate of 1 ml min⁻¹ and eluting with a gradient of 100 mM triethylammonium acetate pH 7.0 (buffer A) and 100 mM triethylammonium acetate in 80% acetonitrile (buffer B). Denaturing polyacrylamide gels were stained with SYBR Gold and visualized by a Typhoon 9400 imaging system.

Reaction of *A. californica* ADPRC with NAD and various alcohols. Reactions were performed on a 30 μl scale in 50 mM HEPES buffer (pH 7) containing 5 mM MgCl₂ at 37 °C for 30 min. NAD (10 mM) was mixed with ADPRC (Sigma-Aldrich, 1.4 μM) and 2 μl of 2-propyn-1-ol, 3-buten-1-ol, 4-pentyn-1-ol or 5-hexyn-1-ol. After the reaction, 2 μl of the reaction mixtures were diluted, filtered and analysed by HPLC. The products were isolated and characterized by electrospray ionization (ESI)-MS and high-resolution MS analysis.

CuAAC. The product of the ADPRC reaction with NAD (10 μl) was isolated by HPLC, freeze-dried, mixed with biotin-PEG₃-azide (Jena Bioscience; three equivalents, 10 mM), and a freshly prepared mixture of CuSO₄ (5 mM) and sodium ascorbate (10 mM) in 30 μl of 50 mM HEPES buffer (pH 7) containing 5 mM MgCl₂, and incubated at 37 °C for 30 min. The products were isolated by HPLC and characterized by ESI-MS and high-resolution MS analysis.

Preparation of model NAD-RNA by *in vitro* transcription. For *in vitro* transcription we used a modification of a previously described method¹³; 0.2 μM of double-stranded DNA (dsDNA) template in transcription buffer (40 mM Tris, 1 mM spermidine, 22 mM MgCl₂, 0.01% Triton-X-100, 5 mM dithiothreitol (DTT), 5% DMSO, pH 8.1), along with 1 mM of each of GTP, ATP, CTP and UTP, with 4 mM NAD (NAD-RNA) or without NAD (unmodified RNA) was mixed in an Eppendorf tube. Then, 5.2 U μl^{-1} of T7 RNA polymerase was added (laboratory-prepared stock) and the transcription mixture was incubated at 37 °C for 4 h. The reaction was stopped by addition of gel loading buffer (10% TBE in formamide containing xylene cyanol and bromophenol blue) and purified over a 12% denaturing polyacrylamide gel using standard electrophoresis conditions (1 \times TBE buffer, run at 25 W for 3 h). Transcription products were excised and eluted in 0.3 M sodium acetate (pH 5.5) overnight at room temperature. The eluted solution was ethanol precipitated and then dissolved in Millipore water.

Application of the capture protocol to model NAD-RNA. The transcribed RNA (NAD-RNA or non-modified RNA, 0.35 nmol) was mixed with 3 μl of 4-pentyn-1-ol and ADPRC (0.4 μM) in 60 μl of 50 mM HEPES and 5 mM MgCl₂ (pH 7) and incubated at 37 °C for 30 min. The RNA was purified by ethanol precipitation, one-third was kept for PAGE analysis and two-thirds were used in the subsequent CuAAC reaction. Biotin-PEG₃-azide (29 μM , calculated as 10 equivalents to RNA) and a freshly prepared mixture of CuSO₄ (1.25 mM), tris(3-hydroxypropyl)triazolylmethylamine (THPTA; 0.625 mM) and sodium ascorbate (2.5 mM) were mixed with purified RNA in 80 μl of 50 mM HEPES and 5 mM MgCl₂ (pH 7) and incubated at 25 °C for 30 min. Half of the sample was then mixed with streptavidin solution (New England Biolabs) at a ratio of 1:1 to added biotin. The samples were mixed with gel loading buffer (10% TBE in formamide containing xylene cyanol and bromophenol blue) and analysed by electrophoresis on a 16% denaturing polyacrylamide gel.

Plasmids and *E. coli* strains. The *nudC* gene was PCR amplified from genomic DNA of *E. coli* K-12 (isolated via GenElute Bacterial Genomic DNA Kit, Sigma-Aldrich). XbaI and NotI restriction sites were introduced during amplification using the primers Fwd NudC and Rev NudC (Supplementary Table 1). The resulting PCR product was digested with XbaI and NotI and cloned into pET-28c vector (Merck Millipore). After Sanger sequencing, the resulting plasmid pET-28c-NudC was transformed into *E. coli* One Shot BL21 Star (DE3) (Life Technologies). The NudC(E178Q) mutant was generated by site-directed mutagenesis using a procedure based on the Phusion Site-Directed Mutagenesis Kit (Thermo Scientific). The resulting plasmid pET-28c-NudC(E178Q) was sequenced and transformed into *E. coli* One Shot BL21 Star (DE3).

For complementation of the *nudC* knockout in *E. coli* JW5548-1 (ref. 31) (*E. coli* Genetic Stock Center (CGSC)), plasmids pTACNudC and pTACNudC(E178Q) that carry *nudC* or *nudC*E178Q were prepared. *nudC* or *nudC*E178Q were amplified from plasmid pET-28c-NudC or pET-28c-NudC(E178Q). EcoRI and SphI restriction sites were introduced during amplification and the PCR product cloned into the pTAC-MAT-2 vector (Sigma). Sequenced plasmids were transformed into *E. coli* JW5548-1.

To create analogous strains that express RNAi, plasmid pHSG 398 (TaKaRa) was subjected to site-directed mutagenesis to obtain the promoter contained in the pUC19 plasmid. pHSG 398-G₋₁-A was then transformed into *E. coli* BW25113 (wild type) (CGSC) *E. coli* JW5548-1 (Δ nudC) and *E. coli* JW5548-1 containing pTACNudC (*nudC* + NudC) or pTACNudC(E178Q) (*nudC* + NudC(E178Q)). All strains used and generated in this work are summarized in Extended Data Fig. 7.

Bacterial RNA isolation. Total RNA from *E. coli* K-12 (strain DSM 498; DSMZ), *E. coli* JM109 (F+ strain; Promega) transformed with pUC19 plasmid, *E. coli* BW25113, *E. coli* JW5548-1 (Δ nudC) and all strains derived thereof was isolated using RNeasy (Molecular Research Center). Bacteria were grown at 37 °C in LB broth until late exponential phase (OD_{600 nm} = 2.0). Expression of NudC in strains containing pTACNudC or pTACNudC(E178Q) was induced with 1 mM isopropyl- β -D-thiogalactoside (IPTG) at OD_{600 nm} = 0.8. To prevent RNA degradation, 40 ml of culture were mixed with 5 ml ice-cold EtOH/phenol stop solution (5% water-saturated phenol, pH < 7.0), and bacteria were harvested by centrifugation at 1,800g at 4 °C for 20 min. The pellet was resuspended in 600 μl lysis solution (2 mg ml⁻¹ in water) and incubated for 5 min at 37 °C. RNeasy (6 ml) was added to the lysate to isolate total RNA according to the supplier's information. In addition, total RNA was DNase I (Thermo Scientific) as well as proteinase K (Sigma Aldrich) digested. DNase I (0.5 U μg^{-1} of total RNA) digestion was performed at 37 °C for 30 min and stopped by the addition of proteinase K (100 μg ml⁻¹), 0.5% (w/v) SDS and 5 mM CaCl₂ at 37 °C for 15 min. Finally, total RNA was phenol/chloroform extracted twice and ethanol precipitated.

Capture protocol for NAD-RNA from total RNA. Three samples of total RNA (100 μg) were treated with 4-pentyn-1-ol (10 μl , minus ADPRC control) or ADPRC (0.85 μM , minus 4-pentyn-1-ol control), or with both 4-pentyn-1-ol and ADPRC (fully treated sample), in 100 μl of 50 mM HEPES and 5 mM MgCl₂ (pH 7) and incubated at 37 °C for 30 min. The reaction was stopped by phenol/ether extraction and RNA was ethanol precipitated. The subsequent CuAAC reaction was performed for all samples on a 100 μl scale in 50 mM HEPES and 5 mM MgCl₂ (pH 7) with biotin-PEG₃-azide (250 μM) and a freshly prepared mixture of CuSO₄ (1 mM), THPTA (0.5 mM) and sodium ascorbate (2 mM) for 30 min at 25 °C. Samples were purified by phenol/ether extraction and ethanol precipitated. Mobio Classic columns were assembled with small filters (MoBioTec, GmbH), and 50 μl of streptavidin immobilized on agarose (CL-4B, Sigma Aldrich) for each sample and were washed three times with 100 μl of immobilization buffer (10 mM HEPES, 1 M NaCl, 5 mM EDTA, pH 7.2). All further washes were performed three times, with the same volume, except for washes with streptavidin washing buffer (five times). The beads were then blocked with 100 μl solution of 100 μg ml⁻¹ BSA (New England Biolabs) in immobilization buffer at room temperature for 20 min at 1,000 r.p.m. After washing the beads, each RNA sample, dissolved in 40 μl immobilization buffer, was incubated with beads at room temperature at 1,000 r.p.m. for 60 min. The beads were washed with streptavidin washing buffer (50 mM Tris-HCl, 8 M urea, pH 7.4) and with 1 \times ligation buffer (50 mM Tris-HCl, 10 mM MgCl₂, 15% dimethylsulphoxide (DMSO), 50 μg ml⁻¹ BSA, pH 7.4). The beads were again blocked as described earlier.

Ligation and amplification protocol. The beads were covered with 20 μl of ligation mixture and incubated at 4 °C for 16 h. The ligation mixture contained adenylated 3'-adaptor³² (5 μM , Adaptor A and Adaptor B; desalted), 50 mM 2-mercaptoethanol, T4 RNA ligase (0.5 U μl^{-1} ; Thermo Scientific), T4 RNA ligase 2, truncated (10 U μl^{-1} ; New England Biolabs) in ligation buffer. The beads were then incubated with 10 μl of 5 M NaCl for 1 h at room temperature at 1,000 r.p.m. (rebinding to streptavidin) and washed first with streptavidin washing buffer, then with 1 \times first strand (FS) buffer (Life Technologies) and again blocked in 100 μl solution of 100 μg ml⁻¹ BSA in 1 \times FS buffer, as described earlier. The beads were then covered with 20 μl of reverse transcription mixture in 1 \times FS buffer containing 0.5 mM dNTPs, BSA (50 μg ml⁻¹), 5 mM DTT and reverse primer (5 μM , RTprimerA or RTprimerB), shaking at 1,000 r.p.m. for 15 min at room temperature. Then SuperScript II Reverse Transcriptase (10 U μl^{-1} ; Life Technologies) was added and the mixture was incubated at 37 °C without shaking for 60 min. The rebinding procedure was repeated (+ 10 μl of 5 M NaCl for 1 h at room temperature shaking at 1,000 r.p.m.) and the beads were washed with streptavidin washing buffer and with 1 \times exonuclease I (Exo I) buffer (Thermo Scientific). The beads were again blocked with 100 μl solution of 100 μg ml⁻¹ BSA in 1 \times Exo I buffer (Thermo Scientific) as described earlier, and again washed with 1 \times Exo I buffer. An exonuclease I digest was performed to remove free single-stranded DNA (ssDNA) primers. Twenty microlitres of 1 \times Exo I buffer with exonuclease I (1 U μl^{-1} ; Thermo Scientific) were incubated with beads for 30 min at 37 °C. After digestion, the beads were again washed with streptavidin washing buffer and with immobilization buffer. In the next step the RNA was digested in 100 μl of 0.15 M NaOH solution without shaking for 25 min at 55 °C. The spin columns were centrifuged and washed with 100 μl of H₂O. The flow-through was ethanol precipitated. The obtained cDNA was dissolved in CTP tailing mixture containing 1.25 mM CTP, 1 \times TdT buffer and terminal deoxynucleotidyl transferase (1 U μl^{-1} ; Thermo Scientific) in a volume of 20 μl , and incubated

for 30 min at 37 °C. The enzyme was thermally denatured by heating to 70 °C for 10 min. A 20 µl mixture for the ligation of the 2nd adaptor (containing 100 mM Tris-HCl, 20 mM MgCl₂, anchor (5 µM, AnchorA annealed with AnchorB), ATP (10 µM) and T4 DNA ligase (1.5 Weiss U µl⁻¹; Thermo Scientific) was added and the solution was incubated for 16 h at 4 °C. The tailed and ligated cDNA was used further in all experiments.

NGS. cDNA (fully treated samples, minus ADPRC controls) was PCR amplified. The regions corresponding to 120–160 bp for sets 1 and 2, or 50–300 bp for set 3 of both minus ADPRC control (A) and fully treated sample (B), were extracted from a 2% agarose gel, purified and either first amplified by PCR, or directly barcoded in a PCR with barcoding primers (Supplementary Table 1). Multiplexed samples were submitted to the CellNetworks Deep Sequencing Core facility for library synthesis using the gDNA kit (New England Biolabs) with TruSeq adaptors from the TruSeq RNA prep kit (Illumina) and sequencing on a HiSeq2000 using the 100 bp single end protocol.

Analysis of NGS data. Using the program RobiNA³³, the barcoded sequencing reads of different samples and controls were demultiplexed. Furthermore, forward and reverse reads were separated according to the adaptor sequence present at the 5'-end of each read. Hereby, we were able to recognize whether a sequencing read was in the original orientation (that is, as on the RNA level) and differentiate between reads that belonged to the 5'- and 3'-ends of those original RNA sequences. This analysis enabled the use of strand-specific feature extraction methods later on in the workflow. During the process, we clipped the barcode and adaptor sequences. We performed all further sequence analysis using the web-based sequence-analysis platform GeneProf³⁴. To this end, the individual barcode- and orientation-separated data sets were uploaded to the platform and the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/): Reverse Complement was used to revert sequences that were in reverse orientation, representing RNA 3'-termini. Furthermore, using the Modify and Filter Sequence tool, we filtered reads according to quality, and removed the additional nucleotides that were added to the cDNA sequence during the tailing reaction, which are thus present as starting Gs at the 5'-termini of forward reads, and, for data set 3, removed adaptor sequences that were present at the 3'-end of certain reads (due to the fact that some PCR products were shorter than the 100 bp read length). A custom reference set for a pUC19-containing *E. coli* strain K-12 was created using reference data from Ensembl³⁵ Bacteria FTP (ASM584v1.16) in which the F-plasmid contained in the JM109 strain was included. We inserted the pUC19 plasmid sequence (GenBank accession L09137.2) manually. For feature annotation, we used the Ensembl Biomart utility.

We aligned all modified forward and reverse reads to the reference data set using the Bowtie³⁶ alignment tool and performed further analysis using the Quantitate Gene Expression tool. For each sample or control, average read numbers as average of normalized read counts (in RPM) from forward and reverse reads were calculated. Enrichment factors were determined as ratio of average read counts of sample B versus control A (Supplementary Data 1–3). For each data set, relative abundance of different reads in sample B was plotted against relative abundance in control A. This analysis revealed that the most abundant reads were clearly separated into an enriched and a non-enriched sequence population. Abundance thresholds $T_{data\ set}$ above which this separation was possible, were determined for each data set separately, as $T_1 = 1,000$, $T_2 = T_3 = 200$. The most abundant sequences were identified as those where the sum of reads in A and B $\geq T_{data\ set}$. Candidate sequences were inspected in the Integrated Genome Browser (IGB)³⁷ and visualized using the GeneProf genome browser. Hits were compared to EcoCyc³⁸ annotations for transcription site annotation and further reference.

Real-time PCR. Quantitative real-time PCR (qPCR) was used to validate the enrichment of sequences identified in the deep sequencing data. All reactions were performed in duplicate in 20 µl scale. Three microlitres cDNA (diluted 1:40 for sRNA and 1:50 for mRNAs) was used as a template for qPCR, which was performed in a Light Cycler 480 instrument (Roche) using the Brilliant III Ultra-Fast SYBR Green qPCR Mastermix (Agilent). Analysis of the data was performed with the Light-Cycler 480 Software. Fully treated sample cDNA experiments were compared with the same experiments performed with minus ADPRC control cDNA and Millipore water was used as a negative control. The primers were designed according to results from deep sequencing analysis for specific target sequences (Supplementary Table 1).

Pull-down of RNAI, GcvB and 5S rRNA. For analysis of endogenous RNAs we used a modification of a previously described method²⁴. Total RNA prepared as described earlier (for LC-MS of RNAI from *E. coli* JM109, OD_{600 nm} = 2.0; for LC-MS of GcvB from *E. coli* K-12, OD_{600 nm} = 1.0) was size-separated by polyacrylamide gel (8%) electrophoresis, and only the region containing RNA of sizes between 75 and 200 nucleotides was eluted and ethanol precipitated. To streptavidin agarose slurry (100 µl, Sigma-Aldrich), washed 2× with 1× PBS buffer (100 µl), biotinylated Hybri-probes (5 nmol) (Supplementary Table 1) were added and incubated at 25 °C, 500 r.p.m. for 10 min. The slurry was washed again two times with 1× PBS

buffer (100 µl), equilibrated two times with pull-down buffer (100 µl, 10 mM Tris-HCl, 0.9 M tetramethylammonium chloride, 0.1 M EDTA, pH 7.8), incubated with 250 µg of RNA (75–200 nucleotides) in pull-down buffer at 65 °C for 10 min and then cooled down to room temperature over 20 min. The slurry was then transferred to Mobicol Classic columns and washed six times with 10 mM Tris-HCl (100 µl, pH 7). The hybridized RNA was released by suspending the slurry in 100 µl of Millipore water, heating of the slurry four times to 75 °C for 10 min, ethanol precipitated, the RNA was further purified over polyacrylamide gel (10%) electrophoresis.

DNase I digestion. RNAI, purified by pull-down, was further treated with DNase I to digest biotinylated hybridization probe. The ethanol-precipitated pull-down RNA (equivalent to 100 µg of 75–200-nucleotide RNA fraction) was dissolved in 20 µl of DNase I buffer containing 0.25 U µl⁻¹ of DNase I. The reaction mixture was incubated at 37 °C for 15 min and the enzyme was thermally denatured by heating to 65 °C for 10 min. RNA was ethanol precipitated before PAGE analysis.

Nuclease P1 digestion. RNAI, GcvB or 5S rRNA pull-down from total RNA was digested by nuclease P1 (0.1 U µl⁻¹; Sigma-Aldrich) in 100 µl of 50 mM ammonium acetate buffer (pH 4.5) at 37 °C for 1 h. The digest was purified over Amicon-Millipore filters 10 kDa (Merck Millipore). The flow-through was concentrated on a Speedvac system to ~20 µl for LC-MS analysis. The nucleotides subjected to LC-MS measurements in positive mode were dephosphorylated with FastAP (Thermo Scientific; 1 U pmol⁻¹ RNA, 1 h at 37 °C) and again purified over Amicon-Millipore filters 3 kDa (Merck Millipore).

HPLC purification of dephosphorylated digest for LC-MS in positive mode. HPLC purification was performed on an Agilent 1200 Series HPLC system equipped with a multi-wavelength detector using a Phenomenex Synergi Fusion-RP column (4 µm, 4.6 × 150 mm) at a flow rate of 0.35 ml min⁻¹ and eluting with a gradient of 5 mM ammonium acetate pH 8.3/acetonitrile. Samples were collected at 17.5–18.5 min and applied to HPLC-MS analysis.

HPLC-MS analysis. HPLC-MS experiments were performed on a Bruker microTOFQ-II ESI mass spectrometer connected to an Agilent 1200 Series HPLC system equipped with a multi-wavelength detector. A Phenomenex Kinetex C18 column (2.6 µm, 2.1 × 100 mm) was used at a flow rate of 0.2 ml min⁻¹ using a gradient of 0.1% aqueous ammonium formate/methanol (LC-MS grade) for nuclease-P1-treated RNA or 100 mM hexafluoroisopropanol/8.6 mM triethylamine pH 8.3 and methanol (LC-MS grade) for the dephosphorylated digest. ESI was used with a capillary voltage of 4,500 V, end plate offset –500 V and collision voltage of 10 eV. The drying gas temperature was 190 °C, the drying gas flow rate was 8 l min⁻¹ and the detector was operated in negative ion mode (RNAI, 5S RNA) or positive ion mode (GcvB, 5S RNA). Analysis of the LC-MS measurements was carried out using Hyphenation Star PP (Version 3.2.44.0) and DataAnalysis (version 4.0, SP 4) software (Bruker Daltonics).

Overexpression and affinity purification of RNase E(1–529), NudC and NudC(E178Q). The RNase E catalytic domain (residues 1–529) (supplied by B. Luisi) was overexpressed and purified as previously described²⁵. IPTG-induced *E. coli* One Shot BL21 Star (DE3) containing the respective plasmid were harvested and lysed by sonication (30 s, 50% power, five times) in HisTrap buffer A (50 mM Tris-HCl pH 7.8, 1 M NaCl, 5 mM MgSO₄, 5 mM 2-mercaptoethanol, 5% glycerol, 5 mM imidazole, 1 tablet per 500 ml complete EDTA-free protease inhibitor cocktail (Roche). The lysate was clarified by centrifugation (37,500g, 30 min, 4 °C) and the supernatant was applied to a Ni-NTA HisTrap column (GE Healthcare). The protein was eluted with a gradient of HisTrap buffer B (HisTrap buffer A with 500 mM imidazole) and analysed by SDS-PAGE. Fractions of interest were concentrated in a micro filter tube (CO 10 kDa, centrifugation at 2,000 r.p.m., 4 °C) and stored in 50 mM Tris-HCl, pH 7.9, 0.5 M NaCl, 50 mM KCl, 1 mM MgCl₂, 5 mM DTT, 50% (v/v) glycerol, 1 tablet per 500 ml complete EDTA-free protease inhibitor cocktail.

E. coli One Shot BL21 Star (DE3) containing pET28c-NudC or pET28c-NudC(E178Q), *E. coli* JW5548-1 pTACNudC or *E. coli* JW5548-1 containing pTACNudC(E178Q) were induced with 1 mM IPTG for 3 h. The harvested cells were lysed by sonication and the His-tagged enzymes were purified similar to RNase E. All purified protein samples were >95% pure, judging from SDS-PAGE.

Preparation of different 5'-terminally labelled RNAs for *in vitro* cleavage assays. Specifically 5'-terminally radiolabelled RNAI, containing either 5'-monophosphate, 5'-triphosphate, 5'-NAD or 5'-NADH was obtained by *in vitro* transcription. As a template for *in vitro* transcription, the sequence of RNAI was amplified from the pUC19 plasmid using the primer pair Fwd T7-φ2.5 RNAI and Rev RNAI (Supplementary Table 1). RNA was transcribed in 200 µl scale in the presence of 40 mM Tris pH 8.1, 1 mM spermidine, 22 mM MgCl₂, 0.01% Triton-X-100, 10 mM DTT, 5% DMSO, 200 mM DNA template, 0.1 mg ml⁻¹ T7 RNA polymerase (lab-prepared stock), 4 mM of CTP, GTP and UTP and 2 mM of ATP. For transcription of monophosphorylated RNAI, the transcription reaction was doped with 200 µCi of radioactive AMP (³²P α-AMP, 6,000 Ci mmol⁻¹; Hartmann Analytic). For preparation of triphosphate-RNAI, 200 µCi of radioactive ATP (³²P γ-ATP, 3,000 Ci mmol⁻¹;

Hartmann Analytic) and for NAD-RNAI 125 μCi of radioactive NAD (adenylate- ^{32}P , 800 Ci mmol^{-1} ; Perkin Elmer) were used. For transcription of NADH-RNAI, ^{32}P -NAD (1.56 μM in 8 mM Tris buffer pH 9.2) was reduced to ^{32}P -NADH by non-enzymatic hydrogen transfer³⁹ in the presence of 50 mM NADH. The formation of 1,4-NADH was verified by a lactate dehydrogenase assay³⁹. ^{32}P -NADH was purified by HPLC. A Phenomenex Kintex C18 column (2.6 μm , 2.1×100 mm) was used and eluted at a flow rate of 1 ml min^{-1} with 0.1 M KH_2PO_4 , 8 mM tetra-*n*-butylammonium hydrogensulphate, pH 6.0, with 0–15% (v/v) acetonitrile. NADH was desalted over a cooled Sep-Pak C18 Cartridge (10 g, 55–105 μm ; Waters) and added to *in vitro* transcription reactions. Transcription reactions were incubated at 37 °C for 4 h and purified on 10% denaturing polyacrylamide gels. Radioactively labelled RNA was eluted from gel slices overnight and isopropanol precipitated. RNA concentrations were determined using the Qubit 2.0 Fluorometer (Life Technologies).

RNase E, RppH and NudC cleavage assays. *In vitro* RNase E and NudC cleavage assays of *in vitro* transcribed monophosphorylated-, triphosphorylated-, NAD- or NADH-RNAI were performed in the presence of 25 mM Tris pH 7.5, 50 mM NaCl, 50 mM KCl, 10 mM MgCl_2 , 1 mM DTT, 5 μM RNA and 1.34 nM RNase E or 4.5 nM NudC/NudC(E178Q). RppH reactions were carried out in presence of 1 \times NEBuffer 2 (New England Biolabs), 5 μM RNA, 0.5 $\text{U } \mu\text{l}^{-1}$ RppH (New England Biolabs). NAD-RNAI and NADH-RNAI cleavage reactions were always performed in the presence of 0.1 $\text{U } \mu\text{l}^{-1}$ FastAP thermosensitive alkaline phosphatase (Thermo Scientific) in order to remove the ^{32}P -phosphate that became accessible only after cleavage of the NAD(H) pyrophosphate bond.

Sequential decapping/digests of 5'-NAD-RNAI or 5'-triphosphate-RNA by NudC/RNase E or RppH/RNase E were carried out as follows: 2.5 μM RNA (^{32}P α -CTP labelled triphosphorylated RNAI or ^{32}P -NAD-RNAI) was treated with NudC or RppH for 20 min at 37 °C, using the conditions described earlier. RNA was separated from enzymes by phenol/chloroform/ether extraction and isopropanol precipitated. RNA pellets were dissolved in water and digested with RNase E using standard conditions described earlier. All time course reactions were performed at 37 °C and reactions were stopped by direct addition of one volume loading dye (10% TBE in formamide containing xylene cyanol and bromophenol blue). Samples were loaded onto 18% denaturing polyacrylamide sequencing gels under standard electrophoresis conditions (1 \times TBE buffer, run at 45 W for 3 h). Gels were visualized using storage phosphor screens (GE Healthcare) and a Typhoon 9400 imager (GE Healthcare) and quantified using ImageQuant software (Molecular Dynamics). Apparent rate constants k_{app} were determined using the 'linear quick fit' tool in OriginPro 8 software (OriginLab).

Quantification of NAD-RNA fractions. The NAD-RNA fraction was quantified in an adapted version of the PABLO assay^{5,40,41}. In the first step, total RNA was dephosphorylated using 0.2 $\text{U } \mu\text{g}^{-1}$ total RNA FastAP thermosensitive alkaline phosphatase (Thermo Scientific). The reaction was performed on a 400 μl scale in the presence of 100 μg total RNA, 1 \times FastAP buffer (Thermo Scientific) for 120 min at 37 °C. RNA was phenol/ether extracted and isopropanol precipitated. Twenty micrograms of the dephosphorylated RNA were incubated with 2 nmol NudC or NudC(E178Q) for 120 min at 37 °C to convert 5'-NAD-capped RNA to 5'-monophosphate RNA (200 μl scale, using NEBuffer 3). Remaining RNA was isopropanol precipitated and redissolved in water. For ligation, 10 μg of the NudC- or NudC(E178Q)-treated total RNA were combined with 5 μM adaptor (Quant-Adaptor; Supplementary Table 1) and 5 μM of the respective splint (Splint-FL-RNAI, Splint-RNAI₋₅ or Splint-6S; Supplementary Table 1) in a final volume of 12 μl , heated to 90 °C for 30 s, and then chilled at room temperature for 5 min. The ligation reaction was incubated for 4 h at 37 °C after addition of 50 mM Tris-HCl, pH 7.5, 0.5 mM DTT, 50 μM ATP, 10 mM MgCl_2 and 60 Weiss U T4 DNA ligase (Thermo Scientific). RNA was isopropanol precipitated, and analysed by northern blotting with target-specific probes for RNAI or 6S RNA. Ligation yields were calculated from the ratio of ligated RNA to ligated plus unligated RNA. Mean values and standard deviations were calculated from yields measured in three independent experiments. Ligation efficiencies were calculated for each individual sample by spiking 1 pmol *in vitro* transcribed NAD ^{32}P -RNAI, or NAD ^{32}P -6S RNA into the respective total RNA sample and treating the mixture with FastAP, NudC (or NudC(E178Q)) and T4 DNA ligase, as described earlier. The major processing product of RNAI, RNAI₋₅, which is generated by RNase E cleavage and thus bears a monophosphate, was not ligated in our NAD quantification assay, as expected. Therefore, to calculate its abundance, in order to estimate the percentage of full-length RNAI, we used a method very similar to the original PABLO assay^{5,40,41}, by performing ligation as described earlier, without prior dephosphorylation or NudC treatment. For calculation of the ligation efficiency of RNAI₋₅, ^{32}P RNAI₋₅ was spiked into total RNA and specifically ligated to the adaptor. For northern blot analysis, RNA was separated on 10% denaturing polyacrylamide gels (7 W, 65 min), transferred by electroblotting to Whatman Nytran SuPerCharge nylon blotting membrane (GE Healthcare) for 2.5 h at 250 mA and UV crosslinked. Membranes

were prehybridized in 15 ml ULTRAhyb Ultrasensitive Hybridization Buffer (Life Technologies) for 2 h at 48 °C. Target-specific RNA riboprobes (Supplementary Table 1) were prepared by *in vitro* transcription as described earlier, incorporating 50 μCi of radioactive CTP (^{32}P α -CTP, 3,000 Ci mmol^{-1} ; Hartmann Analytic) and 50 μCi of radioactive ATP (^{32}P α -ATP, 3,000 Ci mmol^{-1} ; Hartmann Analytic). The *in vitro* transcription product was DNase I (Thermo Scientific) digested and phenol/chloroform extracted. RNA was isopropanol precipitated and the pellet dissolved in 50 μl water. Ten microlitres of the riboprobe were added to the prehybridized membrane and incubated overnight at 48 °C. The blot was washed two times for 5 min with wash solution 1 (2 \times SSC, 0.1% SDS) and two times for 15 min with wash solution 2 (0.25 \times SSC, 0.1% SDS). Results were visualized and quantified as described earlier. For preparation of radioactive RNA size standards, RNA Century Marker Templates (Life Technologies) were *in vitro* transcribed using T7 RNA polymerase in the presence of 40 μCi ^{32}P α -CTP.

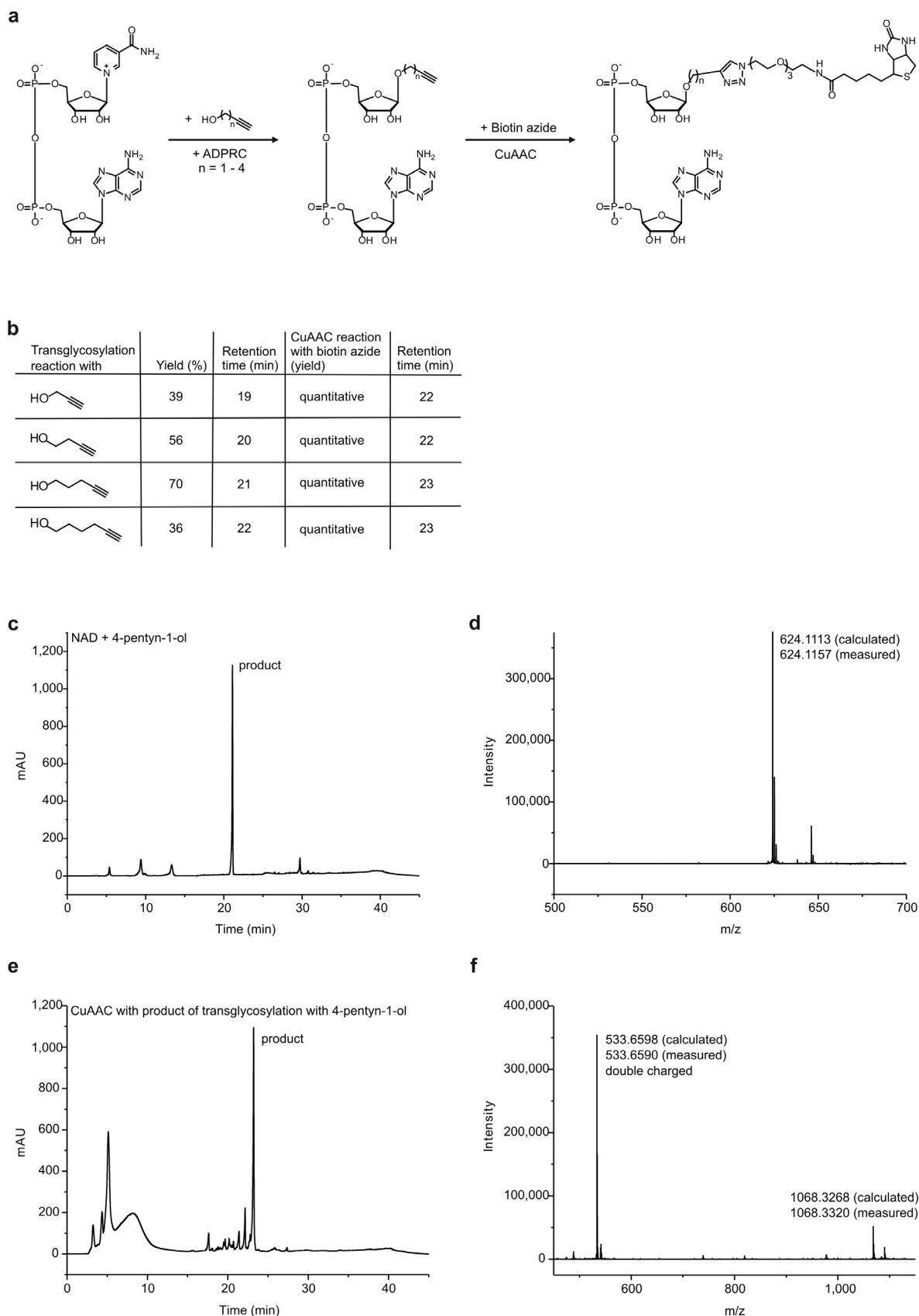
NudC processing of cellular RNAI *in vitro*. One-hundred micrograms total RNA was subjected to alkaline phosphatase FastAP to remove 5'-monophosphate- and 5'-triphosphate termini (400 μl scale, 1 \times FastAP buffer, 0.2 $\text{U } \mu\text{g}^{-1}$ total RNA FastAP thermosensitive alkaline phosphatase for 120 min at 37 °C). RNA was phenol/ether extracted and isopropanol precipitated. One-hundred micrograms dephosphorylated RNA was incubated with 500 pmol NudC or NudC(E178Q). Aliquots were taken at different time intervals, reactions were stopped by flash freezing in liquid nitrogen and the RNA was isopropanol precipitated. NudC processing of RNAI and 6S RNA was analysed by ligation and northern blotting, as described earlier.

RNA half-life determination. Measurements of RNA half-lives were performed in the following strains: wild type and ΔnudC for GcvB or wild type + RNAI and ΔnudC + RNAI for RNAI. Bacteria from overnight cultures were diluted to OD_{600 nm} = 0.1 and grown in LB medium at 37 °C. Half-lives of GcvB and RNAI at OD_{600 nm} = 2.0 were determined by treating cells with rifampicin (500 $\mu\text{g ml}^{-1}$). Samples were taken at time intervals before and after inhibition of the transcription and bacteria from wild type and ΔnudC strain were treated with RNAProtect Bacteria Reagent (Qiagen) to prevent RNA degradation. Then, RNA was isolated and analysed by northern blotting, as described earlier.

- Baba, T. *et al.* Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2**, 2006.0008 (2006).
- Hafner, M. *et al.* Identification of microRNAs and other small regulatory RNAs using cDNA library sequencing. *Methods* **44**, 3–12 (2008).
- Lohse, M. *et al.* RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Res.* **40**, W622–W627 (2012).
- Halbritter, F., Vaidya, H. J. & Tomlinson, S. R. GeneProf: analysis of high-throughput sequencing experiments. *Nature Methods* **9**, 7–8 (2012).
- Flicek, P. *et al.* Ensembl 2013. *Nucleic Acids Res.* **41**, D48–D55 (2013).
- Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
- Nicol, J. W., Helt, G. A., Blanchard, S. G. Jr, Raja, A. & Loraine, A. E. The Integrated Genome Browser: free software for distribution and exploration of genome-scale datasets. *Bioinformatics* **25**, 2730–2731 (2009).
- Keseler, I. M. *et al.* EcoCyc: fusing model organism databases with systems biology. *Nucleic Acids Res.* **41**, D605–D612 (2013).
- Bernofsky, C. & Gallagher, W. J. Nonenzymic hydrogen transfer between reduced and oxidized pyridine nucleotides. *Biochim. Biophys. Acta* **659**, 1–6 (1981).
- Celesnik, H., Deana, A. & Belasco, J. G. PABLO analysis of RNA: 5'-phosphorylation state and 5'-end mapping. *Methods Enzymol.* **447**, 83–98 (2008).
- Celesnik, H., Deana, A. & Belasco, J. G. Initiation of RNA decay in *Escherichia coli* by 5' pyrophosphate removal. *Mol. Cell* **27**, 79–90 (2007).
- Tómcányi, T. & Apirion, D. Processing enzyme ribonuclease E specifically cleaves RNAI. An inhibitor of primer formation in plasmid DNA synthesis. *J. Mol. Biol.* **185**, 713–720 (1985).
- Masukata, H. & Tomizawa, J. Control of primer formation for ColE1 plasmid replication: conformational change of the primer transcript. *Cell* **44**, 125–136 (1986).
- Blomberg, P., Wagner, E. G. & Nordström, K. Control of replication of plasmid R1: the duplex between the antisense RNA, CopA, and its target, CopT, is processed specifically *in vivo* and *in vitro* by RNase III. *EMBO J.* **9**, 2331–2340 (1990).
- Sharma, C. M., Darfeuille, F., Plantinga, T. H. & Vogel, J. A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. *Genes Dev.* **21**, 2804–2817 (2007).
- Thomason, M. K., Fontaine, F., De Lay, N. & Storz, G. A small RNA that regulates motility and biofilm formation in response to changes in nutrient availability in *Escherichia coli*. *Mol. Microbiol.* **84**, 17–35 (2012).
- Jørgensen, M. G. *et al.* Small regulatory RNAs control the multi-cellular adhesive lifestyle of *Escherichia coli*. *Mol. Microbiol.* **84**, 36–50 (2012).
- Jørgensen, M. G., Thomason, M. K., Havelund, J., Valentin-Hansen, P. & Storz, G. Dual function of the McaS small RNA in controlling biofilm formation. *Genes Dev.* **27**, 1132–1145 (2013).
- Rasmussen, A. A. *et al.* A conserved small RNA promotes silencing of the outer membrane protein YbfM. *Mol. Microbiol.* **72**, 566–577 (2009).
- Mandin, P. & Gottesman, S. A genetic approach for finding small RNAs regulators of genes of interest identifies RybC as regulating the DpiA/DpiB two-component system. *Mol. Microbiol.* **72**, 551–565 (2009).

51. Weilbacher, T. *et al.* A novel sRNA component of the carbon storage regulatory system of *Escherichia coli*. *Mol. Microbiol.* **48**, 657–670 (2003).
52. Lease, R. A., Smith, D., McDonough, K. & Belfort, M. The small noncoding DsrA RNA is an acid resistance regulator in *Escherichia coli*. *J. Bacteriol.* **186**, 6179–6185 (2004).
53. Sledjeski, D. D., Gupta, A. & Gottesman, S. The small RNA, DsrA, is essential for the low temperature expression of RpoS during exponential growth in *Escherichia coli*. *EMBO J.* **15**, 3993–4000 (1996).
54. Lease, R. A. & Belfort, M. A trans-acting RNA as a control switch in *Escherichia coli*: DsrA modulates function by forming alternative structures. *Proc. Natl Acad. Sci. USA* **97**, 9919–9924 (2000).
55. Lease, R. A., Cusick, M. E. & Belfort, M. Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. *Proc. Natl Acad. Sci. USA* **95**, 12456–12461 (1998).
56. Göpel, Y., Papenfort, K., Reichenbach, B., Vogel, J. & Görke, B. Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. *Genes Dev.* **27**, 552–564 (2013).
57. Urban, J. H. & Vogel, J. Two seemingly homologous noncoding RNAs act hierarchically to activate *glmS* mRNA translation. *PLoS Biol.* **6**, e64 (2008).
58. Reichenbach, B., Maes, A., Kalamorz, F., Hajnsdorf, E. & Görke, B. The small RNA GlmY acts upstream of the sRNA GlmZ in the activation of *glmS* expression and is subject to regulation by polyadenylation in *Escherichia coli*. *Nucleic Acids Res.* **36**, 2570–2580 (2008).
59. Moon, K. & Gottesman, S. A PhoQ/P-regulated small RNA regulates sensitivity of *Escherichia coli* to antimicrobial peptides. *Mol. Microbiol.* **74**, 1314–1330 (2009).
60. Fozo, E. M. *et al.* Repression of small toxic protein synthesis by the Sib and OhsC small RNAs. *Mol. Microbiol.* **70**, 1076–1093 (2008).
61. Vogel, J. *et al.* RNomics in *Escherichia coli* detects new sRNA species and indicates parallel transcriptional output in bacteria. *Nucleic Acids Res.* **31**, 6435–6443 (2003).
62. Blasi, F. *et al.* Inhibition of transcription of the histidine operon *in vitro* by the first enzyme of the histidine pathway. *Proc. Natl Acad. Sci. USA* **70**, 2692–2696 (1973).
63. Richards, J., Luciano, D. J. & Belasco, J. G. Influence of translation on RppH-dependent mRNA degradation in *Escherichia coli*. *Mol. Microbiol.* **86**, 1063–1072 (2012).
64. Nobelmann, B. & Lengeler, J. W. Molecular analysis of the *gat* genes from *Escherichia coli* and of their roles in galactitol transport and metabolism. *J. Bacteriol.* **178**, 6790–6795 (1996).
65. Tucker, D. L., Tucker, N. & Conway, T. Gene expression profiling of the pH response in *Escherichia coli*. *J. Bacteriol.* **184**, 6551–6558 (2002).
66. Peng, C. A., Oliver, M. J. & Wood, A. J. Is the RehYdrin TrDr3 from *Tortula ruralis* associated with tolerance to cold, salinity, and reduced pH? Physiological evaluation of the TrDr3-orthologue, HdeD from *Escherichia coli* in response to abiotic stress. *Plant Biol.* **7**, 315–320 (2005).
67. Lawther, R. P. & Hatfield, G. W. Multivalent translational control of transcription termination at attenuator of *ilvGEDA* operon of *Escherichia coli* K-12. *Proc. Natl Acad. Sci. USA* **77**, 1862–1866 (1980).
68. Purta, E. *et al.* The *yfhQ* gene of *Escherichia coli* encodes a tRNA: C_m32/U_m32 methyltransferase. *BMC Mol. Biol.* **7**, 23 (2006).
69. Elovson, J. & Vagelos, P. R. Acyl carrier protein. X. Acyl carrier protein synthetase. *J. Biol. Chem.* **243**, 3603–3611 (1968).
70. Lu, P., Vogel, C., Wang, R., Yao, X. & Marcotte, E. M. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nature Biotechnol.* **25**, 117–124 (2007).
71. Rudolph, F. B. & Fromm, H. J. The purification and properties of aspartase from *Escherichia coli*. *Arch. Biochem. Biophys.* **147**, 92–98 (1971).
72. Karsten, W. E. & Viola, R. E. Kinetic studies of L-aspartase from *Escherichia coli*: pH-dependent activity changes. *Arch. Biochem. Biophys.* **287**, 60–67 (1991).
73. Maki, Y., Yoshida, H. & Wada, A. Two proteins, YfiA and YhbH, associated with resting ribosomes in stationary phase *Escherichia coli*. *Genes Cells* **5**, 965–974 (2000).
74. Agafonov, D. E., Kolb, V. A., Nazimov, I. V. & Spirin, A. S. A protein residing at the subunit interface of the bacterial ribosome. *Proc. Natl Acad. Sci. USA* **96**, 12345–12349 (1999).
75. Vila-Sanjurjo, A., Schuwirth, B. S., Hau, C. W. & Cate, J. H. Structural basis for the control of translation initiation during stress. *Nature Struct. Mol. Biol.* **11**, 1054–1059 (2004).
76. Agafonov, D. E., Kolb, V. A. & Spirin, A. S. Ribosome-associated protein that inhibits translation at the aminoacyl-tRNA binding stage. *EMBO Rep.* **2**, 399–402 (2001).
77. Kim, K. S. & Lee, Y. Regulation of 6S RNA biogenesis by switching utilization of both sigma factors and endoribonucleases. *Nucleic Acids Res.* **32**, 6057–6068 (2004).
78. Weber, M. M., French, C. L., Barnes, M. B., Siegle, D. A. & McLean, R. J. A previously uncharacterized gene, *yjiO* (*bsmA*), influences *Escherichia coli* biofilm formation and stress response. *Microbiology* **156**, 139–147 (2010).
79. Robbins, J. C. & Oxender, D. L. Transport systems for alanine, serine, and glycine in *Escherichia coli* K-12. *J. Bacteriol.* **116**, 12–18 (1973).
80. Ghrist, A. C. & Stauffer, G. V. The *Escherichia coli* glycine transport system and its role in the regulation of the glycine cleavage enzyme system. *Microbiology* **141**, 133–140 (1995).
81. Pulvermacher, S. C., Stauffer, L. T. & Stauffer, G. V. Role of the sRNA GcvB in regulation of *cycA* in *Escherichia coli*. *Microbiology* **155**, 106–114 (2009).
82. Hugovieux-Cotte-Pattat, N. & Robert-Baudouy, J. Regulation and transcription direction of *exuR*, a self-regulated repressor in *Escherichia coli* K-12. *J. Mol. Biol.* **156**, 221–228 (1982).
83. Rodionov, D. A., Mironov, A. A., Rakhmaninova, A. B. & Gelfand, M. S. Transcriptional regulation of transport and utilization systems for hexuronides, hexuronates and hexonates in gamma purple bacteria. *Mol. Microbiol.* **38**, 673–683 (2000).
84. Cole, S. T. *et al.* Nucleotide sequence and gene-polypeptide relationships of the *glpABC* operon encoding the anaerobic sn-glycerol-3-phosphate dehydrogenase of *Escherichia coli* K-12. *J. Bacteriol.* **170**, 2448–2456 (1988).
85. Schryvers, A., Lohmeier, E. & Weiner, J. H. Chemical and functional properties of the native and reconstituted forms of the membrane-bound, aerobic glycerol-3-phosphate dehydrogenase of *Escherichia coli*. *J. Biol. Chem.* **253**, 783–788 (1978).
86. Summers, M. C. & Rose, I. A. Proton transfer reactions of methylglyoxal synthase. *J. Am. Chem. Soc.* **99**, 4475–4478 (1977).
87. Plumbridge, J. A. Repression and induction of the *nag* regulon of *Escherichia coli* K-12: the roles of *nagC* and *nagA* in maintenance of the uninduced state. *Mol. Microbiol.* **5**, 2053–2062 (1991).
88. White, R. J. Control of amino sugar metabolism in *Escherichia coli* and isolation of mutants unable to degrade amino sugars. *Biochem. J.* **106**, 847–858 (1968).
89. Freundlich, M., Ramani, N., Mathew, E., Sirko, A. & Tsui, P. The role of integration host factor in gene expression in *Escherichia coli*. *Mol. Microbiol.* **6**, 2557–2563 (1992).
90. Nelson, W. C., Howard, M. T., Sherman, J. A. & Matson, S. W. The *traY* gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid *oriT*. *J. Biol. Chem.* **270**, 28374–28380 (1995).
91. Dhavan, G. M., Crothers, D. M., Chance, M. R. & Brenowitz, M. Concerted binding and bending of DNA by *Escherichia coli* integration host factor. *J. Mol. Biol.* **315**, 1027–1037 (2002).
92. Yanagisawa, T., Sumida, T., Ishii, R., Takemoto, C. & Yokoyama, S. A paralog of lysyl-tRNA synthetase aminoacylates a conserved lysine residue in translation elongation factor P. *Nature Struct. Mol. Biol.* **17**, 1136–1143 (2010).
93. Kaiser, K. & Murray, N. E. Physical characterisation of the “Rac prophage” in *E. coli* K12. *Mol. Gen. Genet.* **175**, 159–174 (1979).
94. Caldas, T. *et al.* The FtsJ/RrmJ heat shock protein of *Escherichia coli* is a 23 S ribosomal RNA methyltransferase. *J. Biol. Chem.* **275**, 16414–16419 (2000).
95. Theobald, A., Springer, M., Grunberg-Manago, M., Ebel, J. P. & Gieger, R. Tertiary structure of *Escherichia coli* tRNA^{Thr3} in solution and interaction of this tRNA with the cognate threonyl-tRNA synthetase. *Eur. J. Biochem.* **175**, 511–524 (1988).
96. Serres, M. H. *et al.* A functional update of the *Escherichia coli* K-12 genome. *Genome Biol.* **2**, RESEARCH0035 (2001).
97. Tschowri, N., Busse, S. & Hengge, R. The BLUF-EAL protein YcgF acts as a direct anti-repressor in a blue-light response of *Escherichia coli*. *Genes Dev.* **23**, 522–534 (2009).
98. Daley, D. O. *et al.* Global topology analysis of the *Escherichia coli* inner membrane proteome. *Science* **308**, 1321–1323 (2005).
99. Lhoest, J. & Colson, C. Cold-sensitive ribosome assembly in an *Escherichia coli* mutant lacking a single methyl group in ribosomal protein L3. *Eur. J. Biochem.* **121**, 33–37 (1981).
100. Heurgué-Hamard, V., Champ, S., Engström, A., Ehrenberg, M. & Buckingham, R. H. The *hemK* gene in *Escherichia coli* encodes the N⁵-glutamine methyltransferase that modifies peptide release factors. *EMBO J.* **21**, 769–778 (2002).
101. White, P. J., Millar, G. & Coggin, J. R. The overexpression, purification and complete amino acid sequence of chorismate synthase from *Escherichia coli* K12 and its comparison with the enzyme from *Neurospora crassa*. *Biochem. J.* **251**, 313–322 (1988).
102. Imamura, R. *et al.* Identification of the *cpdA* gene encoding cyclic 3',5'-adenosine monophosphate phosphodiesterase in *Escherichia coli*. *J. Biol. Chem.* **271**, 25423–25429 (1996).
103. Long, C. W. & Pardee, A. B. Cytidine triphosphate synthetase of *Escherichia coli* B. I. Purification and kinetics. *J. Biol. Chem.* **242**, 4715–4721 (1967).
104. Py, B., Higgins, C. F., Krisch, H. M. & Carposis, A. J. A DEAD-box RNA helicase in the *Escherichia coli* RNA degradosome. *Nature* **381**, 169–172 (1996).
105. Jeanningros, R., Creuzet-Sigal, N., Frixon, C. & Cattaneo, J. Purification and properties of a debanching enzyme from *Escherichia coli*. *Biochim. Biophys. Acta* **438**, 186–199 (1976).
106. Ballicora, M. A., Iglesias, A. A. & Preiss, J. ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev.* **67**, 213–225 (2003).
107. Joly, J. C. & Wickner, W. The SecA and SecY subunits of translocase are the nearest neighbors of a translocating preprotein, shielding it from phospholipids. *EMBO J.* **12**, 255–263 (1993).
108. Wada, A. & Sako, T. Primary structures of and genes for new ribosomal proteins A and B in *Escherichia coli*. *J. Biochem.* **101**, 817–820 (1987).
109. Sparrow, C. P. & Raetz, C. R. Purification and properties of the membrane-bound CDP-diglyceride synthetase from *Escherichia coli*. *J. Biol. Chem.* **260**, 12084–12091 (1985).
110. Langley, K. E. & Kennedy, E. P. Partial purification and properties of CTP:phosphatidic acid cytidyltransferase from membranes of *Escherichia coli*. *J. Bacteriol.* **136**, 85–95 (1978).
111. Kanehara, K., Ito, K. & Akiyama, Y. YaeL (EcfE) activates the sigma(E) pathway of stress response through a site-2 cleavage of anti-sigma(E), RseA. *Genes Dev.* **16**, 2147–2155 (2002).
112. Alba, B. M., Leeds, J. A., Onufryk, C., Lu, C. Z. & Gross, C. A. DegS and YaeL participate sequentially in the cleavage of RseA to activate the sigma(E)-dependent extracytoplasmic stress response. *Genes Dev.* **16**, 2156–2168 (2002).
113. Takeda, K., Akimoto, C. & Kawamukai, M. Effects of the *Escherichia coli* *sfsA* gene on *mal* genes expression and a DNA binding activity of SfsA. *Biosci. Biotechnol. Biochem.* **65**, 213–217 (2001).

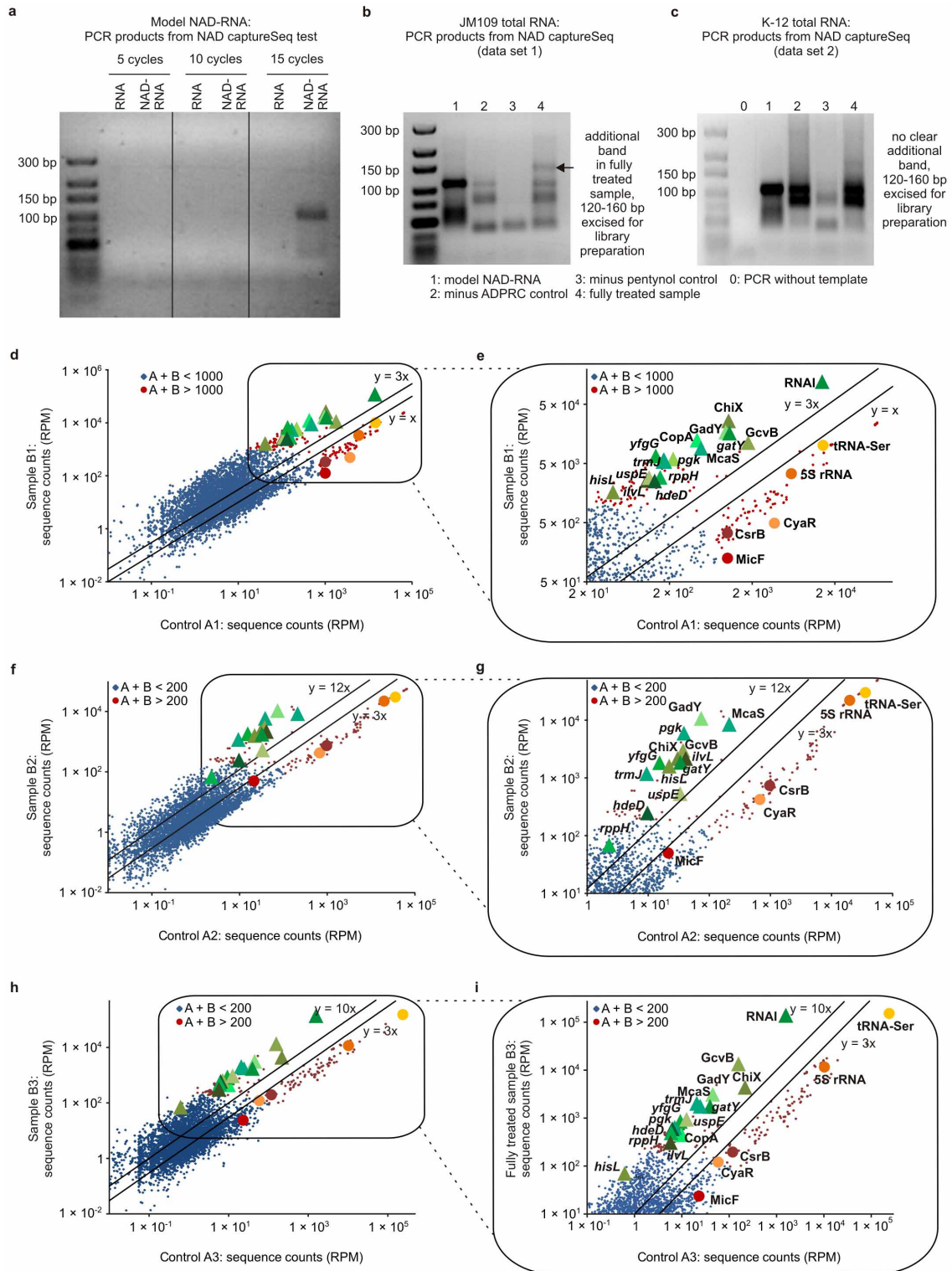
114. Kawamukai, M. *et al.* Nucleotide sequence and characterization of the *sfs1* gene: *sfs1* is involved in CRP*-dependent *mal* gene expression in *Escherichia coli*. *J. Bacteriol.* **173**, 2644–2648 (1991).
115. Paul, B. J. *et al.* DksA: a critical component of the transcription initiation machinery that potentiates the regulation of rRNA promoters by ppGpp and the initiating NTP. *Cell* **118**, 311–322 (2004).
116. Clark, R. B. & Ogilvie, J. W. Aspartokinase I-homoserine dehydrogenase I of *Escherichia coli* K12. Subunit molecular weight and nicotinamide-adenine dinucleotide phosphate binding. *Biochemistry* **11**, 1278–1282 (1972).
117. Starnes, W. L. *et al.* Threonine-sensitive aspartokinase-homoserine dehydrogenase complex, amino acid composition, molecular weight, and subunit composition of the complex. *Biochemistry* **11**, 677–687 (1972).
118. Burr, B., Walker, J., Truffa-Bachi, P. & Cohen, G. N. Homoserine kinase from *Escherichia coli* K12. *Eur. J. Biochem.* **62**, 519–526 (1976).
119. Johansen, J., Eriksen, M., Kallipolitis, B. & Valentin-Hansen, P. Down-regulation of outer membrane proteins by noncoding RNAs: unraveling the cAMP-CRP- and σ E-dependent CyaR-ompX regulatory case. *J. Mol. Biol.* **383**, 1–9 (2008).
120. De Lay, N. & Gottesman, S. The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. *J. Bacteriol.* **191**, 461–476 (2009).
121. Andersen, J., Forst, S. A., Zhao, K., Inouye, M. & Delihas, N. The function of *micF* RNA. *micF* RNA is a major factor in the thermal regulation of OmpF protein in *Escherichia coli*. *J. Biol. Chem.* **264**, 17961–17970 (1989).
122. Deighan, P., Free, A. & Dorman, C. J. A role for the *Escherichia coli* H-NS-like protein StpA in OmpF porin expression through modulation of *micF* RNA stability. *Mol. Microbiol.* **38**, 126–139 (2000).
123. Trotochaud, A. E. & Wassarman, K. M. 6S RNA function enhances long-term cell survival. *J. Bacteriol.* **186**, 4978–4985 (2004).
124. Wassarman, K. M. & Storz, G. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* **101**, 613–623 (2000).
125. Kim, E. Y., Shin, M. S., Rhee, J. H. & Choy, H. E. Factors influencing preferential utilization of RNA polymerase containing sigma-38 in stationary-phase gene expression in *Escherichia coli*. *J. Microbiol.* **42**, 103–110 (2004).
126. Muto, A. *et al.* Structure and function of 10Sa RNA: *trans*-translation system. *Biochimie* **78**, 985–991 (1996).
127. Roche, E. D. & Sauer, R. T. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J.* **18**, 4579–4589 (1999).
128. Kawano, M., Reynolds, A. A., Miranda-Rios, J. & Storz, G. Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* **33**, 1040–1050 (2005).



Extended Data Figure 1 | Development of the NAD capture protocol.

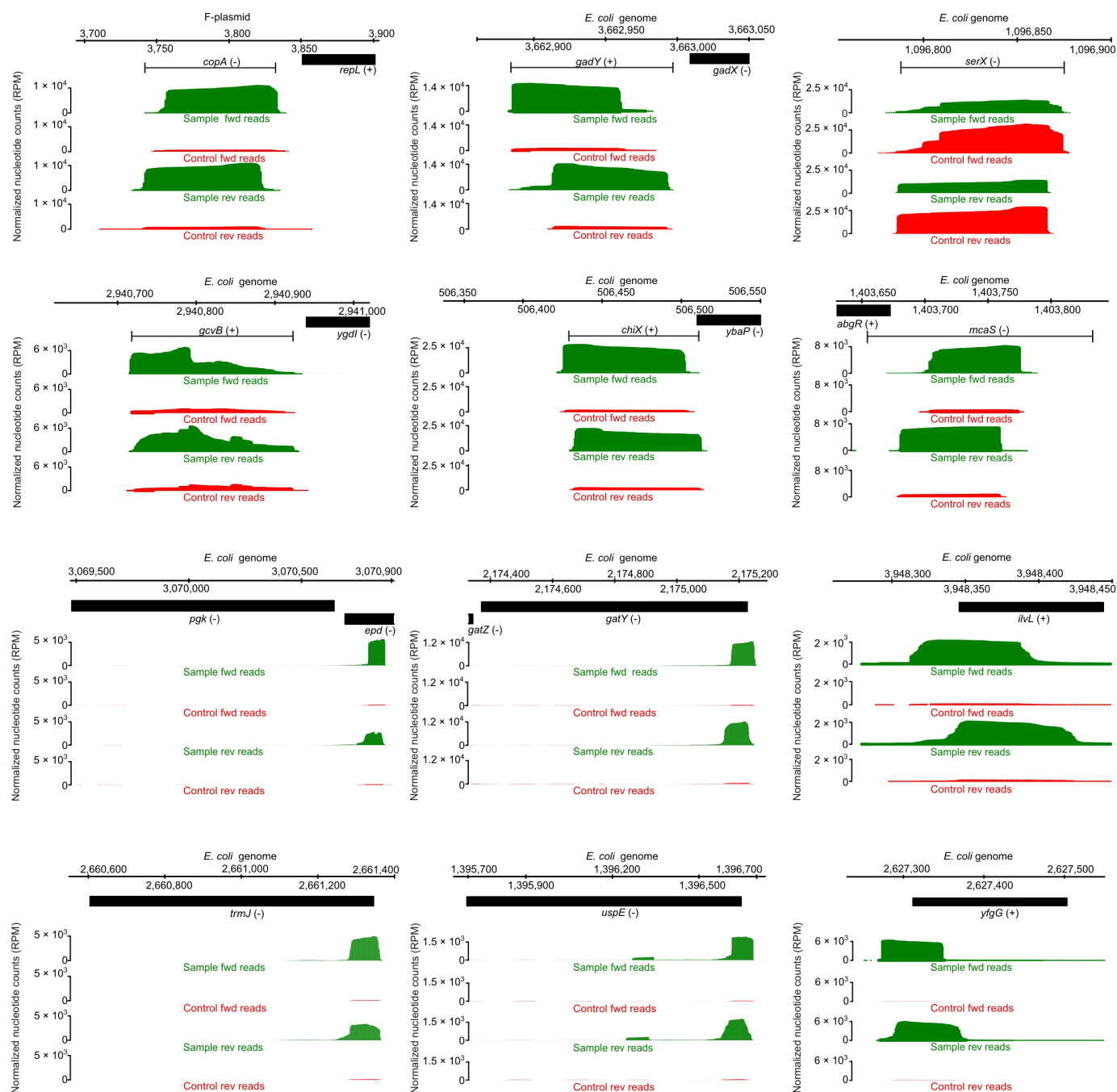
a, General scheme: ADPRC-catalysed transglycosylation with different alcohols, followed by 'click' biotinylation with biotin-PEG₃-azide. **b**, Isolated yields and HPLC retention times. **c**, HPLC chromatogram of the ADPRC-catalysed reaction of NAD with 4-pentyn-1-ol. **d**, High-resolution ESI mass

spectrum of the transglycosylation product. **e**, HPLC chromatogram of the copper-catalysed azide-alkyne cycloaddition (CuAAC) reaction mixture with the product from **c**. **f**, High-resolution ESI mass spectrum of the CuAAC reaction product.



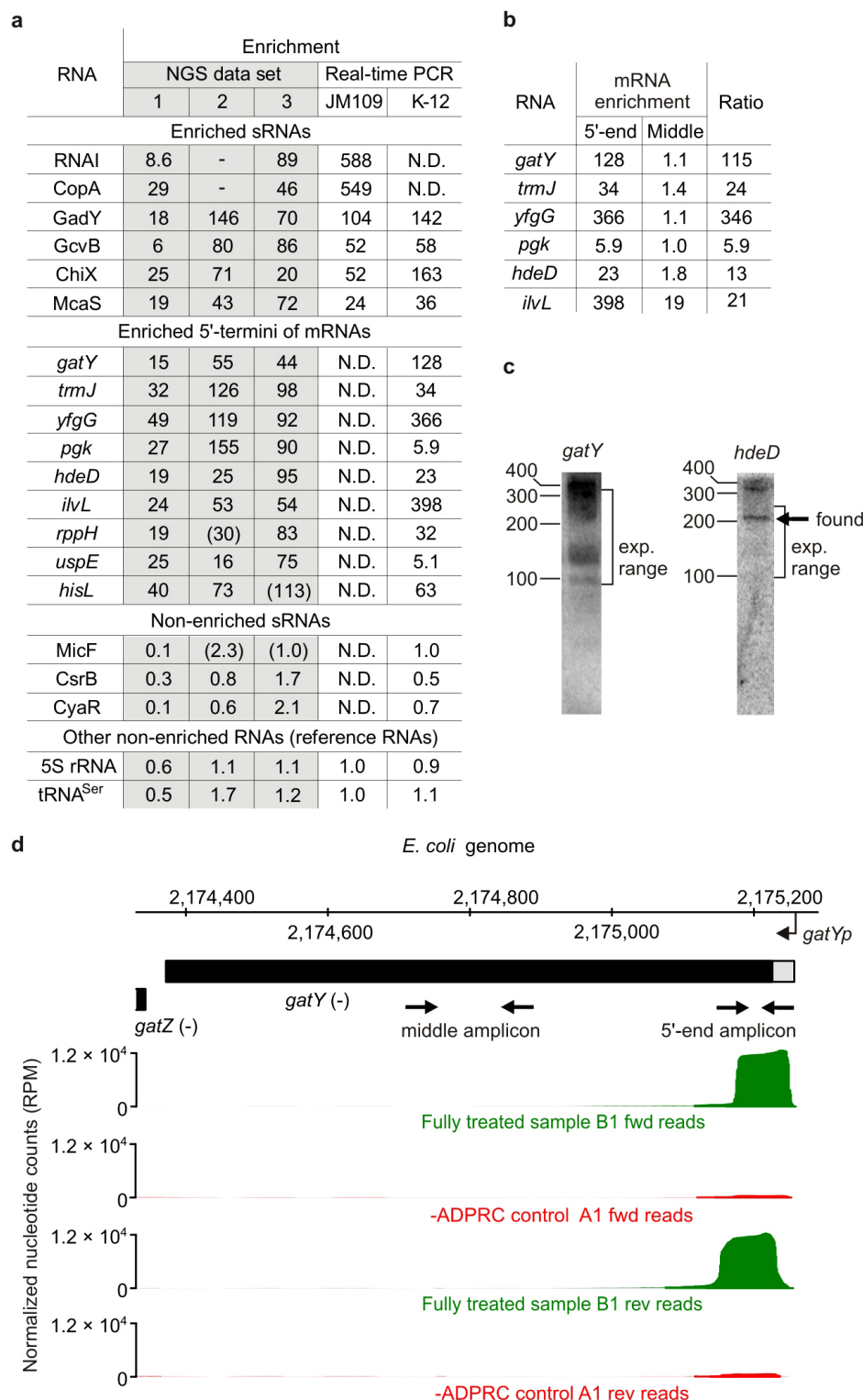
Extended Data Figure 2 | NAD capture protocol applied to RNA and comparison of the enrichment in NGS data sets from three different NAD captureSeq experiments. **a–c**, Agarose gels of the PCR amplification products of cDNA obtained from the NAD capture protocol, applied to *in vitro* transcribed unmodified RNA and *in vitro* transcribed model NAD-RNA after 5 (left), 10 (middle), and 15 (right) cycles (**a**); and total RNA from *E. coli* JM109 (**b**) or K-12 (**c**). The region (120–160 bp) around the band present in the JM109 sample (~150 bp, arrow) that occurred only in the fully treated sample but not in the controls was chosen for NGS libraries of minus ADPRC control A1 and fully treated sample B1, in agreement with the expected size (<200 nucleotides) of NAD-modified RNA according to previous observations². The intense band of ~150 bp size was later attributed to the

most abundant enriched RNA, RNAI, which was absent in the K-12 sample. For this reason, the K-12 sample did not contain such a prominent additional band. For comparability, the same region (120–160 bp) was excised for library preparation for K-12 control A2 and sample B2. To achieve a more comprehensive picture and study the influence of sample preparation on data, a third library was prepared from JM109 (control A3, sample B3) choosing a broader size range (50–300 bp). **d–i**, Abundance of RNAs in fully treated sample (B) versus minus ADPRC control (A) in data set 1 (JM109, 120–160 bp; **d**, **e**), data set 2 (K-12, 120–160 bp; **f**, **g**), and data set 3 (JM109, 50–300 bp; **h**, **i**) found by strand-specific NGS analysis (RPM, average of forward and reverse read counts; Supplementary Data 1–3). Triangles, enriched; circles, non-enriched.



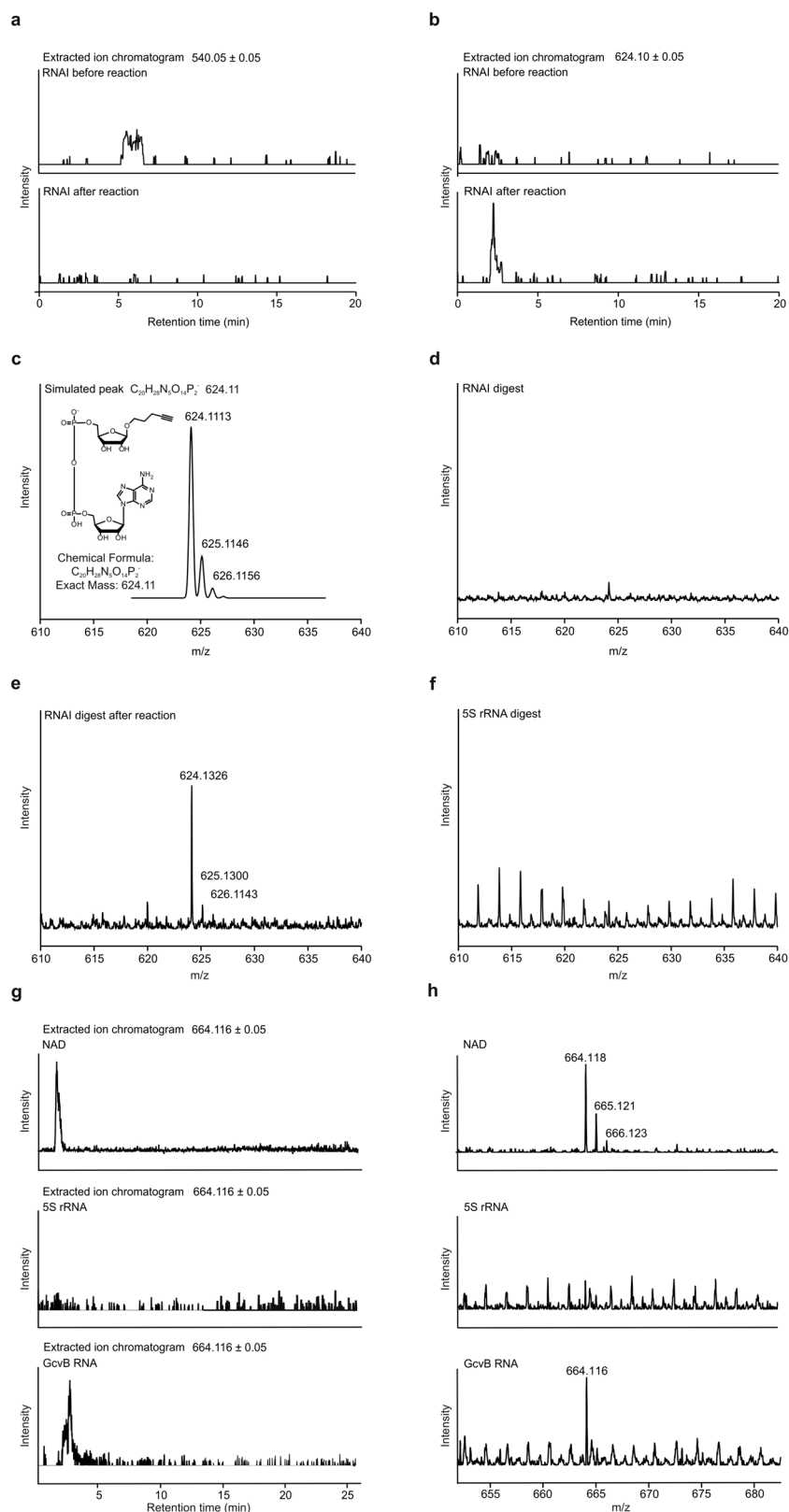
Extended Data Figure 3 | NGS analysis of specific sequences found in control A1 and sample B1 in forward and reverse reads. Analysis of sRNAs, sRNA-like 5'-terminal fragments of mRNAs, and non-enriched tRNA^{Ser}. Normalized nucleotide counts derived from NGS reads (JM109, 120–160 bp) of control A1 (red) and fully treated sample B1 (green) are plotted for the regions

on the *E. coli* K-12 genome or the respective plasmid in which the sequences are situated. A black bar indicates the protein-coding sequence of mRNAs, while non-coding RNAs are indicated by a thin black line. Plus and minus signs designate the orientation of genes in the genome/plasmid.



Extended Data Figure 4 | NGS, real-time PCR, and northern blot analysis of RNAs enriched in NAD captureSeq. **a**, Enrichment of selected sequences calculated from NGS read numbers or real-time PCR data. Numbers in parentheses show enrichment of genes below abundance threshold *T* in the respective data set. N.D., not determined. **b**, Real-time PCR analysis of selected mRNAs for 5'-termini (regions found to be enriched) and regions in the middle of the mRNA. Data show that the 5'-termini are highly enriched, while middle regions are not, or are considerably less enriched. All real-time PCR experiments were performed in duplicate for the same cDNA, and no-template

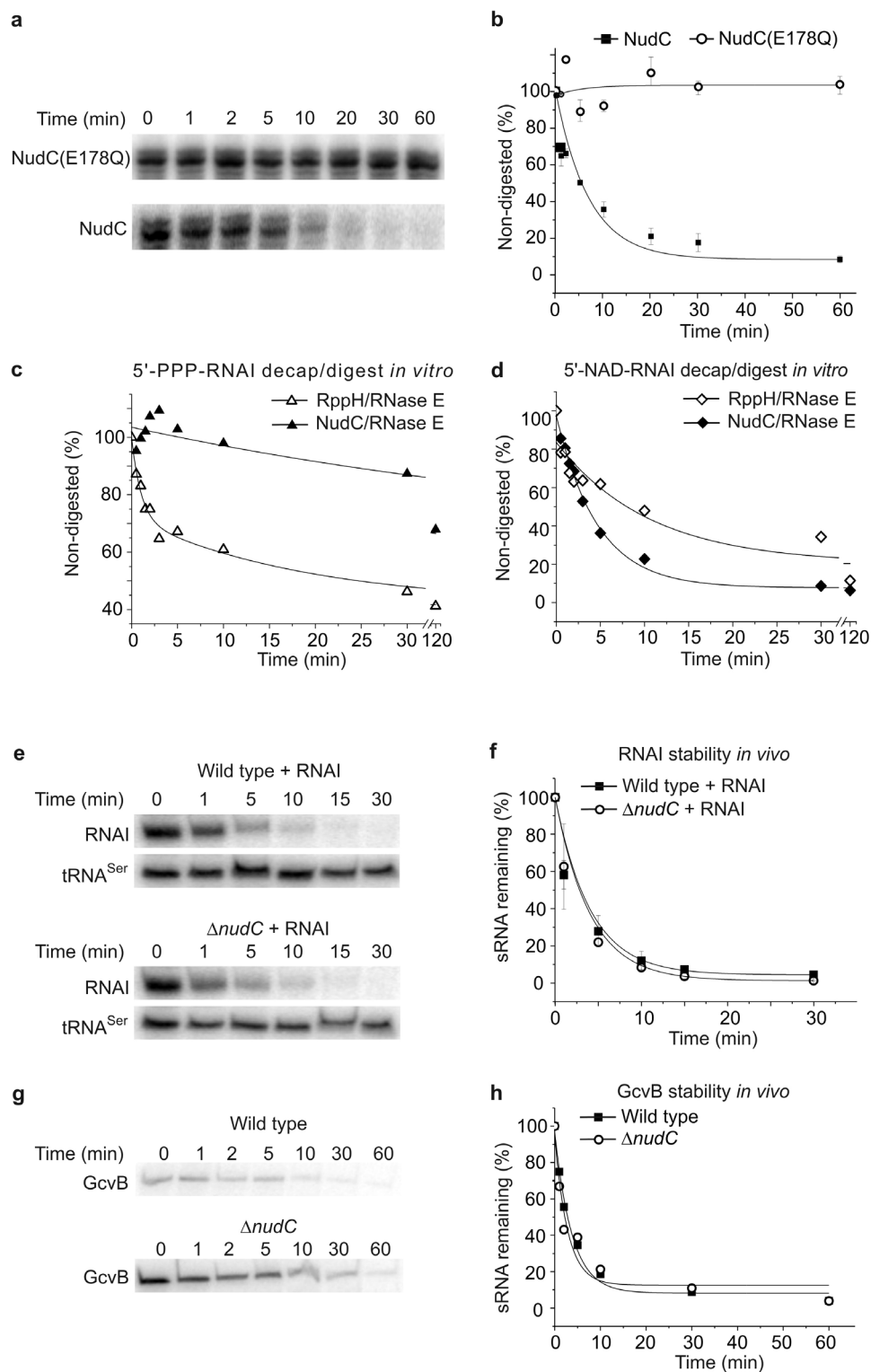
controls were used to assess primer dimerization. **c**, Northern blot analysis. The expected size range (~100 to 340 nucleotides for *gatY*, ~100 to 240 nucleotides for *hdeD*) is indicated and for *hdeD* the fragment of expected size is indicated by an arrow. **d**, Example of an enriched 5'-terminus, *gatY*. The 5'-untranslated region is indicated in grey, the coding sequence in black. Minus sign indicates that the gene is encoded on the minus-strand of the genome, thus in reverse orientation. The position of the promoter (*gatYp*) is indicated by an arrow. Regions amplified in real-time PCR are indicated by opposing arrows.



Extended Data Figure 5 | Mass-spectrometric analysis of the ADPRC-catalysed transglycosylation reaction of NAD-RNAI from *E. coli* total RNA and 4-pentyn-1-ol, and analysis of the 5'-modification status of GcvB.

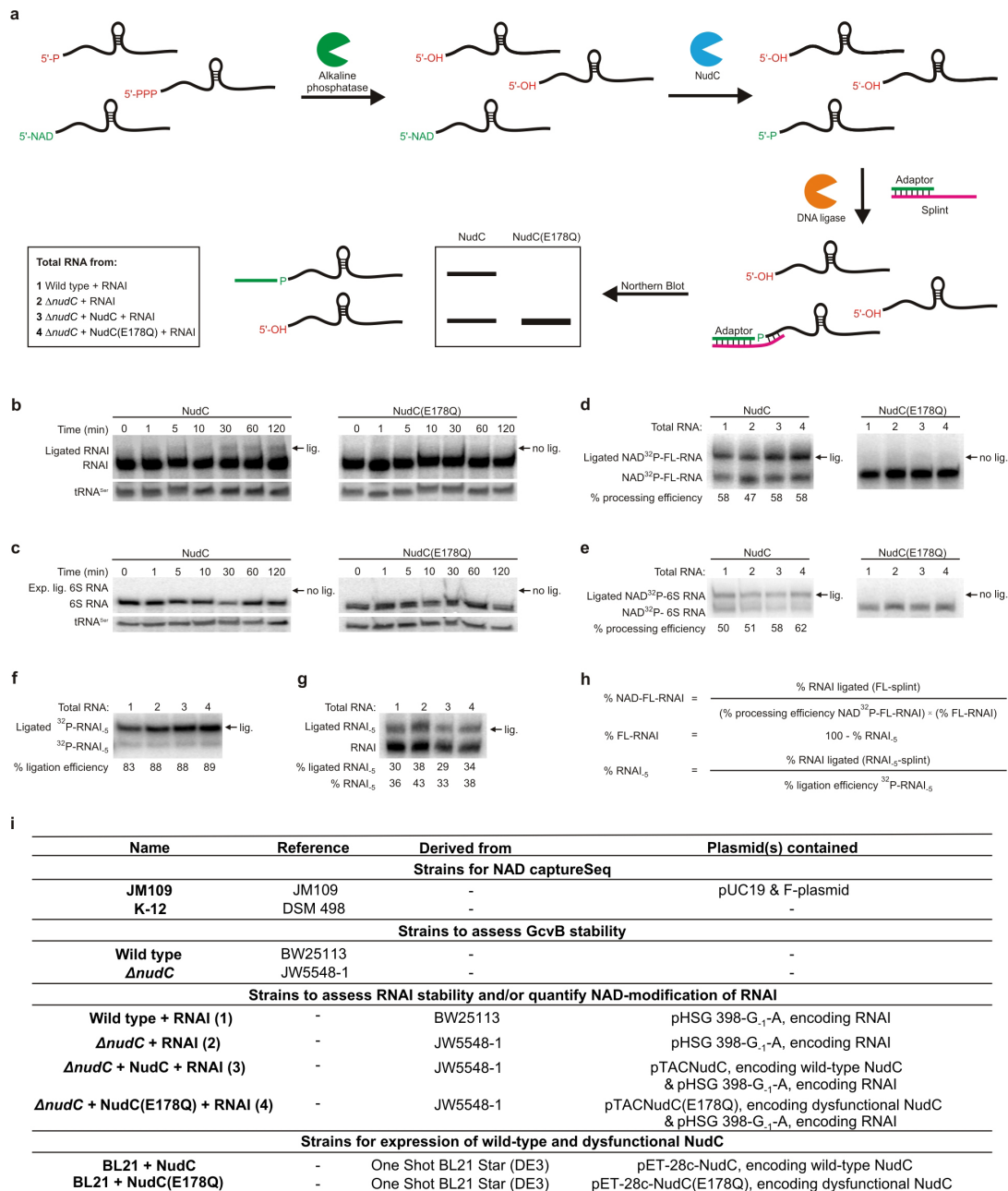
a, b, Disappearance of the ADP-ribose peak ($[M-H]^-$ $m/z = 540.05$, elution at 5–6 min) (**a**), and appearance of a peak at $m/z = 624.10$ (elution time 2.0–2.5 min) upon ADPRC reaction (**b**). Samples were digested with nuclease P1 before mass spectrometry. **c**, Simulated signal pattern of the expected reaction product (isotopic distribution). **d–f**, Observed mass spectrum in the

chromatography fractions eluting at 2–2.5 min of untreated RNAI (**d**), RNAI treated with ADPRC and 4-pentyn-1-ol and then digested by nuclease P1 (**e**) and 5S rRNA treated with ADPRC and 4-pentyn-1-ol and then digested by nuclease P1 (**f**). **g**, HPLC-MS extracted ion chromatograms of the peak $[M]^+$ $m/z = 664.116$ of free NAD, and nuclease-P1-digested 5S rRNA and GcvB RNA. **h**, High-resolution mass spectra of the HPLC fractions eluting at 2.2–2.4 min.



Extended Data Figure 6 | NudC processing, 5'-dependent sensitization to RNase E processing by RppH or NudC, and investigation of RNA stability *in vivo* depending on NudC expression. **a, b**, Decapping activity on *in vitro* transcribed NAD-RNAI of wild-type NudC enzyme and the E178Q point mutant. Gel-electrophoretic assay, monitoring the disappearance of a site-specifically installed radiolabel (see Fig. 3c). Mean of a technical duplicate \pm s.d.

is shown. **c, d**, Successive decapping and digestion of NAD-RNAI and triphosphate-RNAI. Time course of RppH/RNase E or NudC/RNase E decapping/digests of triphosphate-RNAI (**c**) or NAD-RNAI (**d**). **e–h**, Determination of RNAI (mean of biological duplicate \pm s.d.) (**e, f**) and GcvB (**g, h**) cellular half-life depending on NudC expression by rifampicin stop assay. Detection by northern blotting, tRNA^{Ser} as loading control.



Extended Data Figure 7 | Determination of the 5'-modification status of cellular RNAs by a modified PABLO assay^{5,40}.

a, Schematic overview of NAD quantification by dephosphorylation, NudC cleavage and splinted ligation. Total RNA is isolated from different *E. coli* strains: wild type + RNAI (1); $\Delta nudC$ + RNAI (2); $\Delta nudC$ + NudC + RNAI (3); or $\Delta nudC$ + NudC(E178Q) + RNAI (4) (for strain details, see **i**). To remove all 5'-monophosphate and 5'-triphosphate termini, total RNA samples are subjected to alkaline phosphatase treatment. Then, to remove 5'-NAD and produce 5'-monophosphate termini, the dephosphorylated samples are treated with NudC (or NudC(E178Q) as a negative control). Finally, the resulting 5'-monophosphate RNA is ligated specifically to an adaptor by splinted ligation. Ligated fractions are determined by northern blotting. **b**, **c**, Validation of time-dependent *in vitro* NudC processing of cellular sRNAs (RNAI, 6S RNA) in total RNA from *E. coli* wild type + RNAI (1), using the assay described in **a**. Aliquots were taken after different times of NudC treatment and subjected to ligation as shown. tRNA^{Ser} is blotted as loading control. **d**, **e**, Quantification of the efficiency of NudC treatment and splinted ligation (processing efficiency) for NAD-full-length RNAI (NAD-FL-RNAI) and NAD-6S RNA. To calculate the percentage of NAD modification (shown in Fig. 4), it is necessary to normalize to the cumulative processing efficiency of NudC treatment and splinted ligation of NAD-RNA, which is determined experimentally for each total RNA

preparation. To this end, *in vitro* transcribed radiolabelled NAD³²P-FL-RNAI (**d**) or NAD³²P-6S RNA (**e**) is spiked into total RNAs isolated from the four *E. coli* strains mentioned above. The mixture is then subjected to the treatment described in **a**. Processing efficiencies of radiolabelled RNAs are quantified radiometrically after gel electrophoresis and electroblotting. **f**, **g**, Quantification of the percentage of RNAI₋₅. Cellular RNAI consists of a mixture of full-length and processed species, among them RNAI₋₅ (ref. 42), which arises from RNase E cleavage, thus contains a 5'-monophosphate, and lacks only five 5'-terminal nucleotides. Both species exhibit very similar electrophoretic mobility. Here, we were interested only in the percentage of NAD-modified FL-RNAI. Thus, we quantified the percentage of RNAI₋₅ by PABLO analysis (splinted ligation of non-treated total RNA with RNAI₋₅-specific adaptor). Carrying out the respective control (*in vitro* transcribed monophosphate³²P-RNAI₋₅, spiked into total RNA samples) (**f**), we determined the percentage of ligation product, and normalized to ligation efficiencies to obtain the percentage of RNAI₋₅ (**g**). With the help of the values determined in **d**–**g**, we then determined the percentage of FL-RNAI and the percentage of NAD-FL-RNAI (shown in Fig. 4a), using the formulas shown in **h**. Values shown in this figure represent replicate 1. (Expected) ligation products are marked with arrows (**b**–**g**). **i**, Strains used in this study.

Extended Data Table 1 | Listing of all enriched high-abundance sequences from three different NAD captureSeq experiments

RNA	Data set			Comments (e.g., function of sRNA or encoded protein)
	1	2	3	
Enriched sRNAs (with real-time PCR validation)				
RNAI	+	0	+	Only in JM109; copy control RNA of pUC19 ^{17,43} . (AS)
CopA	+	0	+	Only in JM109; copy control RNA of F-plasmid ^{16,44} ; RNase III mediated regulation. (AS)
GadY	+	+	+	Involved in acid stress response ^{15,21} , regulates stability and RNase III degradation ¹⁵ of transcripts. (AS)
GcvB	+	+	+	Involved in acid stress response ²⁰ , regulates ABC transporters on translational level ⁴⁵ . (AS)
McaS	+	+	+	[Found as IS061 (non-strand-specific), due to unclear annotation.]Involved in biofilms. Regulates translation and binds CsrA ⁴⁶⁻⁴⁸ . (AS)
ChiX	+	+	+	Involved in acid stress response ²¹ , regulates porin ChiP ⁴⁹ and DpiA/DpiB two-component system ⁵⁰ . (AS)
Enriched sRNAs (without real-time PCR validation)				
CsrC	+	(+)	+	Binds to and regulates protein CsrA ⁵¹ .
DsrA	+	(+)	+	Acid stress response ⁵² , low-temperature expression of RpoS ⁵³ ; regulates target stability ⁵⁴ & translation ⁵⁵ . (AS)
GlmY	(+)	+	+	(Tke1); binds to RapZ ⁵⁶ protein, thereby regulates GlmZ, which regulates GlmS ^{57,58} .
MgR	+	+	+	(STnc560); induced at low Mg ²⁺ concentration; involved in resistance against polymyxin B ⁵⁹ .
RNAII	+	0	+	Pattern suggests RNAII could be enriched due to pairing with RNAI. Replication primer in RNAI-RNAII system of pUC19 ⁴³ . (AS)
SibC	-	+	+	(QUAD) Repression of small toxic protein lbcC ⁶⁰ . (AS)
SibD	+	+	+	(QUAD) Repression of small toxic protein lbcD ⁶⁰ . (AS)
SibE	+	+	+	(QUAD) Repression of small toxic protein lbcE ⁶⁰ . (AS)
SraC	+	(+)	+	(RyeA); involved in acid stress response ²¹ ; RNase III induced cleavage of RyeB ⁶¹ . (AS)
Enriched 5'-termini (with real-time PCR validation)				
hisL	+	+	(+)	His leader peptide; attenuates transcription of <i>hisLGDCBHAFI</i> operon (enzymes for histidine biosynthesis ⁶²).
rppH	+	(+)	+	RNA pyrophosphohydrolase; removes pyrophosphate, targets RNA for RNase E degradation ^{5,63} .
pgk	+	+	+	Mapped to <i>epd</i> 3'-terminus. Matches <i>pgk</i> transcript annotation (EcoCyc); phosphoglycerate kinase ¹⁸ .
gatY	+	+	+	Tagatose-1,6-bisphosphate aldolase; part of the galactitol catabolism pathway ⁶⁴ .
hdeD	+	+	+	Acid resistance membrane protein ^{19,65,66} .
ilvL	+	+	+	Ilv leader peptide; attenuates transcription of <i>ilvLGMEDA</i> operon (enzymes for isoleucine and valine biosynthesis ⁶⁷).
yfgG	+	+	+	Predicted protein.
trmJ	+	+	+	tRNA U _m 32 methyltransferase/tRNA C _m 32 methyltransferase ⁶⁸ .
uspE	+	+	+	Universal stress protein E. Implication in UV-stress, also in other stresses ²² .
Enriched 5'-termini (without real-time PCR validation)				
acpP	+	+	+	Subunit of holo-acyl carrier protein ⁶⁹ ; important role in fatty acid biosynthesis; one of the most abundant proteins in <i>E. coli</i> ⁷⁰ .
aspA	+	+	+	Aspartate ammonia lyase ⁷¹ ; catalyses the amination of fumarate to yield aspartate in pH-dependent manner ⁷² .
raiA	+	+	+	Stationary phase translation inhibitor and ribosome stability factor; association with 70S ribosome in stationary phase ⁷³ ; stabilization of 70S ribosome ⁷⁴ ; inhibits translation initiation ⁷⁵ and elongation ⁷⁶ in response to stress.
ssrS	+	+	+	Mapped to <i>zapA</i> 3'-end. Matches EcoCyc annotation of longer 6S transcript, processed by RNase E to mature 6S RNA ⁷⁷ .
bsmA	+	(+)	+	Involved in biofilm stress and motility. Mutants are less resistant to acid and peroxide stress ⁷⁸ .
cycA	+	(+)	+	Serine/alanine/glycine/cycloserine: H ⁺ symporter ^{79,80} ; regulated by GcvB on translation level ⁸¹ .
exuR	+	(+)	+	DNA-binding transcription repressor ^{82,83} of synthesis of exuR, galacturonate & glucuronate catabolism & transport genes.
glpA	+	(+)	+	Glycerol-3-phosphate dehydrogenase (anaerobic), large subunit; binds FAD ⁸⁴ .
glpD	+	(+)	+	Glycerol-3-phosphate dehydrogenase (aerobic); binds FAD ⁸⁵ .
mgsA	(+)	+	+	Methylglyoxal synthase ⁸⁶ .
nagC	+	+	+	Mapped to <i>nagA</i> 3'-end; NagC: DNA-binding dual transcriptional regulator of aminosugar biosynthesis ^{87,88} .
ihfA	(+)	+	+	Mapped to <i>pheT</i> 3'-end; IHF: integration host factor (α subunit); IHF important in gene expression ⁸⁹ and DNA bending ^{90,91} .
poxA	+	(+)	+	Does not match annotated transcription start (5' of it); EF-P-lysine lysyltransferase ⁹² ; aminoacylates EF-P at Lys34.
racR	+	(+)	+	Rac prophage ⁹³ ; predicted DNA-binding transcriptional regulator ³⁸ .
rlmE	+	+	(+)	23S rRNA 2'-O-ribose U2552 methyltransferase ⁹⁴ .
thrS	+	(+)	+	Threonyl-tRNA synthetase ⁹⁵ .
traY	+	0	+	Only in JM109: protein encoded on F-plasmid; promotes nicking at <i>oriT</i> ⁹⁰ .
ucpA	+	(+)	+	Predicted oxidoreductase; predicted to bind NAD ⁹⁶ .
ycgE	+	(+)	+	BluR DNA-binding transcriptional repressor; represses genes involved in biofilm formation & acid resistance ⁹⁷ .
yghB	+	+	(+)	Conserved inner membrane protein ⁹⁸ .
Enriched sequences with unclear annotation (without real-time PCR validation)				
aroC?	+	+	+	Mapped to <i>yfcB</i> 3'-end (protein glutamine methyltransferase ^{99,100}); AroC: chorismate synthase; binds flavin ¹⁰¹ ; no transcript annotation.
cpdA?	+	(+)	+	Mapped to <i>yqiB</i> 3'-end (predicted dehydrogenase ³⁸); CpdA: cAMP phosphodiesterase ¹⁰² ; no transcript annotated in EcoCyc.
eno?	+	+	+	Mapped to <i>pyrG</i> 3'-end (CTP synthase ¹⁰³); does not match EcoCyc transcript annotation of <i>eno</i> . Enolase, part of degradosome ¹⁰⁴ .
glgC?	+	(+)	+	Mapped to <i>glgX</i> 3'-end (debranching enzyme ¹⁰⁵); does not match transcript annotation of <i>glgC</i> . ADP-Glc PPase ¹⁰⁶ .
rpmJ?	+	+	+	Mapped to <i>secY</i> 3'-end (involved in protein secretion ¹⁰⁷); no transcript annotation for <i>rpmJ</i> . 50S ribosomal subunit protein L36 ¹⁰⁸ .
rseP?	+	(+)	+	Mapped to <i>cdsA</i> 3'-end (CDP-diglyceride synthetase ^{109,110}); no match: transcript annotation of <i>rseP</i> . signal peptide peptidase ^{111,112} .
sfsA?	+	+	+	Starts 3'- of <i>sfsA</i> 5'-end (predicted DNA-binding transcriptional regulator of maltose metabolism ^{113,114}); does not match transcript annotation of <i>dkSA</i> . RNA polymerase-binding transcription factor ¹¹⁵ .
thrB?	+	(+)	+	Mapped to <i>thrA</i> 3'-end (NAD(P)H-dependent kinase/dehydrogenase ^{116,117}); no transcript annotation for <i>thrB</i> . Homoserine kinase ¹¹⁸ .
yfcO?	+	+	(+)	Mapped to <i>yfcO</i> 3'-end (predicted protein, unknown function ³⁸); could be 5'-end of <i>yfcB</i> (see <i>aroC?</i>).
Examples of non-enriched genes (with real-time PCR validation)				
5S	-	-	-	5S ribosomal RNA.
tRNA ^{Ser}	-	-	-	Serine tRNA.
CsrB	-	-	-	Involved in acid stress response ²¹ ; binds to and regulates CsrA ⁵¹ .
CyaR	-	-	(-)	Down-regulates e.g., outer membrane protein OmpX ¹¹⁹ and NAD-synthetase NadE ¹²⁰ by regulating translation. (AS)
MicF	-	(-)	(-)	Down-regulates porin OmpF by inhibiting translation and destabilizing transcript in response to diverse stresses ^{121,122} . (AS)
Examples of non-enriched genes (without real-time PCR)				
16S	-	-	-	16S ribosomal RNA.
23S	-	-	-	23S ribosomal RNA.
6S	-	-	-	Mature 6S; regulates transcription (stat. phase), repressing transcription from certain σ 70- and σ S-dependent promoters ^{77,123-125} .
tmRNA	-	-	-	Transfer-messenger RNA; helps releasing stalled ribosomes from mRNA without stop-codon or in lack of appropriate tRNA ^{126,127} .
RyfD	-	-	-	Mapped to <i>clpB</i> . Matches transcription start annotated on EcoCyc ⁴⁴ . An abundant non-coding sRNA of unknown function ¹²⁸ .

Enriched sequences were found to be enriched in at least two out of three data sets. +, Sequence is enriched and more than the abundance threshold T; (+), sequence is enriched but less than T; -, sequence is not enriched and more than T; (-), sequence is less than T and not enriched; 0, sequence not found in data set (not expressed in strain); (AS), sRNA acts via antisense mechanism. Names in parentheses are alternative names, as found in the annotation used for GeneProf analysis. If sequences were mapped to a different gene than found or suspected to be the true gene (often the case for 5'-ends), the gene to which it was mapped is given in Comments. References 43–128 are cited in this table.

ERRATUM

doi:10.1038/nature14220

Erratum: Antarctic glaciation caused ocean circulation changes at the Eocene–Oligocene transition

A. Goldner, N. Herold & M. Huber

Nature **511**, 574–577 (2014); doi:10.1038/nature13597

In this Letter, the Fig. 2a legend should state ‘Ocean surface temperature anomalies’ and the Fig. 2b legend should state ‘Zonally averaged ocean temperature anomaly’, rather than vice versa. This has been corrected in the online versions of the paper.

ERRATUM

doi:10.1038/nature14271

Erratum: Human intracellular ISG15 prevents interferon- α/β over-amplification and auto-inflammation

Xianqin Zhang, Dusan Bogunovic, Béatrice Payelle-Brogard, Véronique Francois-Newton, Scott D. Speer, Chao Yuan, Stefano Volpi, Zhi Li, Ozden Sanal, Davood Mansouri, Ilhan Tezcan, Gillian I. Rice, Chunyuan Chen, Nahal Mansouri, Seyed Alireza Mahdavian, Yuval Itan, Bertrand Boisson, Satoshi Okada, Lu Zeng, Xing Wang, Hui Jiang, Wenqiang Liu, Tiantian Han, Delin Liu, Tao Ma, Bo Wang, Mugen Liu, Jing-Yu Liu, Qing K. Wang, Dilek Yalnizoglu, Lilliana Radoshevich, Gilles Uzé, Philippe Gros, Flore Rozenberg, Shen-Ying Zhang, Emmanuelle Jouanguy, Jacinta Bustamante, Adolfo García-Sastre, Laurent Abel, Pierre Lebon, Luigi D. Notarangelo, Yanick J. Crow, Stéphanie Boisson-Dupuis, Jean-Laurent Casanova & Sandra Pellegrini

Nature **517**, 89–93 (2015); doi:10.1038/nature13801

In this Letter, author J.-L.C. was inadvertently missing one of their affiliations; they should also have been associated with affiliation number 2 (St. Giles Laboratory of Human Genetics of Infectious Diseases, Rockefeller Branch, The Rockefeller University, New York, New York 10065, USA). This has been corrected in the online versions of the paper.